



Prévisions d'ensemble à l'échelle saisonnière : mise en place d'une dynamique stochastique

Lauriane Saunier-Batté

► To cite this version:

Lauriane Saunier-Batté. Prévisions d'ensemble à l'échelle saisonnière : mise en place d'une dynamique stochastique. Sciences de la Terre. Université Paris-Est, 2013. Français. NNT : 2013PEST1001 . pastel-00795478

HAL Id: pastel-00795478

<https://pastel.archives-ouvertes.fr/pastel-00795478>

Submitted on 28 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Thèse présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité : Sciences et Techniques de l'Environnement

par

Lauriane SAUNIER BATTÉ

Ecole Doctorale : SCIENCES, INGÉNIERIE ET ENVIRONNEMENT

*Prévisions d'ensemble à l'échelle saisonnière :
mise en place d'une dynamique stochastique*

Thèse soutenue le 23 janvier 2013 devant le jury composé de :

Francisco Doblas-Reyes	<i>Rapporteur</i>
Laurent Terray	<i>Rapporteur</i>
Fabio D'Andrea	<i>Examineur</i>
Nicholas Hall	<i>Examineur</i>
Marc Bocquet	<i>Directeur de thèse</i>
Michel Déqué	<i>CoDirecteur de thèse</i>

A Gabrielle, ma jolie petite puce, qui a su ajouter la juste dose d'aléatoire dans mon quotidien de doctorante pendant presque deux ans.

REMERCIEMENTS

Ecrire des remerciements, c'est avant tout reconnaître que sans soutiens, guides et conseils, je n'aurais pas pu écrire cette thèse... Dans les moments de doute, les périodes de flottement, j'ai toujours pu trouver quelqu'un qui par une parole ou un geste m'a transmis l'impulsion nécessaire pour rebondir. Je ne saurai citer toutes ces personnes, et je leur prie donc par avance de m'en excuser.

Je souhaite tout particulièrement remercier Michel, qui a su distiller ses remarques avisées au cours de ces trois ans et quelques mois, en me laissant la liberté d'explorer des pistes par moi-même, tout en me remettant sur les rails lorsque c'était nécessaire. Merci aussi à Marc d'avoir endossé la responsabilité de directeur de thèse, et d'avoir su répondre avec efficacité et disponibilité aux contraintes administratives.

J'aimerais également exprimer ma gratitude envers Paco Doblas-Reyes et Laurent Terray, qui ont accepté de lire et évaluer mes travaux, malgré des emplois du temps très contraignants. Merci également à Nick Hall et Fabio D'Andrea d'avoir participé au jury de thèse, et de m'avoir suggéré des pistes intéressantes à explorer pour la suite.

Ces trois années n'auraient pas été aussi agréables sans la bonne humeur qui règne dans le Groupe de Modélisation Grande Echelle et Climat. Merci à tous de m'avoir accueillie et intégrée dans cet environnement a priori intimidant. Je souhaite tout particulièrement saluer quelques personnes qui auront marqué ces trois ans : Jeanne pour son oreille attentive et les PhD Comics, Anne-So pour son dynamisme et sa bonne humeur infaillible, Boutheina, Gaëlle et Agathe pour leurs cœurs en or, John-Py pour ses blagounettes en toutes circonstances, Fanny pour les fous rires, Milou pour son côté détonnant, Pierre pour son côté rassurant, Yannick pour sa sérénité, Julien pour ses vanes bien senties. Merci également à Benoît, Vincent, Cloclo, Sam, Aurore, Marine, Virginie, Guillaume, Aurélien, Hervé G., Silvana, Bertrand, Gilles, Fabrice, David, Hervé D., Audrey, Marania et j'en oublie, pour les bons moments passés ensemble. Merci à mes co-lofleurs puis à Jeff d'avoir supporté mon côté autiste-avec-boules-quiès pendant cette thèse ! Matthieu Che, je ne t'oublie pas : super-binôme depuis la 4A, tu as toujours été une béquille sur laquelle m'appuyer dans les moments difficiles. :) Mille mercis !

Dans les moments de doute, ou simplement pour me changer les idées, j'ai pu trouver avec mes coéquipières de l'Astro, chez mes amis et ma famille un bol d'air frais. Le déroulement et l'organisation de cette thèse ont été quelque peu chamboulés par l'arrivée de Gabrielle, petite bouille d'ange au caractère bien trempé. Merci à toi ma puce d'avoir été sage et compréhensive quand ta maman était un peu débordée. Et surtout, merci Loïc, pour ton amour, ton soutien et ton aide inconditionnels, sans lesquels tout ceci n'aurait jamais été possible.

SOMMAIRE

Résumé	vii
Abstract	ix
Avant-Propos	1
I Introduction aux prévisions saisonnières numériques	3
1 Prévisions saisonnières dynamiques : fondements et état de l’art	5
1.1 Fondements théoriques de la prévision à longue échéance	5
1.1.1 Que peut-on prévoir à l’échelle saisonnière ?	6
1.1.2 Modes de variabilité à l’échelle saisonnière	8
1.1.3 Autres composantes du système climatique à l’échelle saisonnière	14
1.2 De la compréhension de l’ENSO à la modélisation numérique du climat	17
1.2.1 Prévisions statistiques à l’échelle saisonnière	17
1.2.2 L’apparition des modèles numériques de climat	18
1.2.3 La prévision saisonnière numérique	18
1.2.4 Techniques de correction des modèles	20
1.3 Prise en compte des incertitudes dans les modèles dynamiques	22
1.3.1 Incertitudes liées aux conditions initiales	23
1.3.2 Incertitudes liées aux choix de modélisation	24
1.3.3 Perturbations stochastiques des modèles de climat	26
1.4 Questions abordées dans cette thèse	30
<i>Chapter Summary</i>	33
2 Outils et modèles	35
2.1 Composantes du modèle couplé CNRM-CM5.1	35
2.1.1 Le modèle d’atmosphère ARPEGE-Climat	35
2.1.2 Les composantes océan, glace de mer et surfaces continentales	39
2.1.3 Le couplage des différentes composantes : assemblage du modèle de climat	41
2.1.4 Implémentation d’une prévision saisonnière	41
2.2 Evaluation des prévisions	45
2.2.1 A quoi compare-t-on les prévisions ?	45
2.2.2 Les scores d’évaluation des prévisions déterministes	48
2.2.3 Scores probabilistes	50
2.3 Outils statistiques	53
2.3.1 La technique de nudging, ou guidage des modèles	54

2.3.2	Analyse en composantes principales et applications	54
2.3.3	Classification automatique	56
2.4	Synthèse	59
	<i>Chapter Summary</i>	60
II	L'approche multi-modèle	61
3	Prévisions d'ensemble des précipitations sur l'Afrique	63
3.1	Rétro-prévisions du projet ENSEMBLES des précipitations africaines	64
3.1.1	Résumé de l'article publié dans Tellus, Series A	64
3.1.2	Article : Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models : skill of the ENSEMBLES project multi-model ensemble forecasts	65
3.2	Prévisibilité sur le Golfe de Guinée et le Sahel, et liens avec les TSO	89
3.2.1	Le Golfe de Guinée et le Sahel : deux régions contrastées	89
3.2.2	Liens entre précipitations sur l'Afrique de l'ouest et TSO, et représentation dans les modèles	90
3.2.3	Impact du délai de prévision : performances des prévisions pour la saison JAS	93
3.3	Prévisions du modèle CNRM-CM5.1	95
3.3.1	Prévision de la mousson africaine	95
3.3.2	Le cas des autres régions d'Afrique	101
3.4	Synthèse	103
	<i>Chapter Summary</i>	105
III	La dynamique stochastique	107
4	Perturbations stochastiques de la dynamique du modèle ARPEGE-Climat	109
4.1	Théorie et méthodologie retenue	110
4.1.1	De la théorie à l'implémentation dans le modèle	110
4.1.2	Erreurs de tendance initiale du modèle ARPEGE-Climat	111
4.2	Perturbations dérivées par nudging d'anomalie	112
4.2.1	Motivations et mise en œuvre	112
4.2.2	Caractéristiques des perturbations	114
4.2.3	Simulations de la saison d'hiver (DJF)	116
4.2.4	Simulations de la saison d'été (JJA)	120
4.2.5	Discussion	121
4.3	Perturbations dérivées par nudging itéré	121
4.3.1	Motivations et mise en œuvre	121
4.3.2	Dynamique stochastique aléatoire	127
4.3.3	Impact des corrections moyennes	134
4.3.4	Quelle fréquence des perturbations ?	135
4.4	Synthèse	140
	<i>Chapter Summary</i>	142
5	Vers un conditionnement des perturbations	143
5.1	Quelles améliorations possibles ?	144

5.1.1	Résumé de l'article publié dans Geophysical Research Letters	144
5.1.2	Article : A stochastic method for improving seasonal predictions	145
5.1.3	Cas des perturbations issues du nudging d'anomalie	151
5.1.4	Discussion	158
5.2	Conditionnement en fonction des températures de surface de l'océan	158
5.2.1	Principe de base	158
5.2.2	Résultats obtenus	159
5.3	Classification par la fonction de courant	165
5.3.1	Principe et mise en œuvre	165
5.3.2	Fonction de courant à 500 hPa	165
5.3.3	Fonction de courant à 200 hPa	172
5.3.4	Interprétation des résultats	176
5.4	Synthèse	180
	<i>Chapter Summary</i>	182
Conclusion		183
6 Conclusions et Perspectives		185
6.1	Synthèse générale et discussion	185
6.1.1	Prévisions des précipitations africaines : état de l'art des modèles	185
6.1.2	Dynamique stochastique aléatoire	186
6.1.3	Conditionnement des perturbations	189
6.1.4	Principales limites de l'étude	190
6.2	Perspectives	190
6.2.1	Dans un futur proche	191
6.2.2	A plus long terme	192
Liste des Figures		195
Liste des Tableaux		201
Acronymes		207
Bibliographie		211

RÉSUMÉ

La prévision d'ensemble à l'échelle saisonnière avec des modèles de circulation générale a connu un essor certain au cours des vingt dernières années avec la croissance exponentielle des capacités de calcul, l'amélioration de la résolution des modèles, et l'introduction progressive dans ceux-ci des différentes composantes (océan, atmosphère, surfaces continentales et glace de mer) régissant l'évolution du climat à cette échelle. Malgré ces efforts, prévoir la température et les précipitations de la saison à venir reste délicat, non seulement sur les latitudes tempérées mais aussi sur des régions sujettes à des aléas climatiques forts comme l'Afrique de l'ouest pendant la saison de mousson. L'une des clés d'une bonne prévision est la prise en compte des incertitudes liées à la formulation des modèles (résolution, paramétrisations, approximations et erreurs). Une méthode éprouvée est l'approche multi-modèle consistant à regrouper les membres de plusieurs modèles couplés en un seul ensemble de grande taille. Cette approche a été mise en œuvre notamment dans le cadre du projet européen ENSEMBLES, et nous montrons qu'elle permet généralement d'améliorer les rétro-prévisions saisonnières des précipitations sur plusieurs régions d'Afrique par rapport aux modèles pris individuellement.

On se propose dans le cadre de cette thèse d'étudier une autre piste de prise en compte des incertitudes du modèle couplé CNRM-CM5, consistant à ajouter des perturbations stochastiques de la dynamique du modèle d'atmosphère ARPEGE-Climat. Cette méthode, baptisée « dynamique stochastique », consiste à introduire des perturbations additives de température, humidité spécifique et vortacité corrigeant des estimations d'erreur de tendance initiale du modèle. Dans cette thèse, deux méthodes d'estimation des erreurs de tendance initiale ont été étudiées, basées sur la méthode de nudging (guidage) du modèle vers des données de référence. Elles donnent des résultats contrastés en termes de scores des rétro-prévisions selon les régions étudiées.

Si on estime les corrections d'erreur de tendance initiale par une méthode de nudging itéré du modèle couplé vers les réanalyses ERA-Interim, on améliore significativement les scores sur l'hémisphère Nord en hiver en perturbant les prévisions saisonnières en tirant aléatoirement parmi ces corrections. Cette amélioration est accompagnée d'une nette réduction des biais de la hauteur de géopotential à 500 hPa. Une rétro-prévision en utilisant des perturbations dites « optimales » correspondant aux corrections d'erreurs de tendance initiale du mois en cours de prévision montre l'existence d'une information à l'échelle mensuelle qui pourrait permettre de considérablement améliorer les prévisions. La dernière partie de cette thèse explore l'idée d'un conditionnement des perturbations en fonction de l'état du modèle en cours de prévision, afin de se rapprocher si possible des améliorations obtenues avec ces perturbations optimales.

ABSTRACT

Over the last twenty years, research in ensemble predictions at a seasonal timescale using general circulation models has undergone a considerable development due to an exponential growth rate of computing capacities, improved model resolution and the introduction of more and more components (ocean, atmosphere, land surface and sea-ice) that have an impact on climate at this time scale. Regardless of these efforts, predicting temperature and precipitation for the upcoming season is a difficult task, not only over mid-latitudes but also over regions subject to high climate risk, like West Africa during the monsoon season. One key to improving predictions is to represent model uncertainties (due to resolution, parametrizations, approximations and model error). The multimodel approach is a well-tried method which consists in pooling members from different individual coupled models into a single superensemble. This approach was undertaken as part of the European Commission funded ENSEMBLES project, and we find that it usually improves seasonal precipitation re-forecasts over several regions of Africa with respect to individual model predictions.

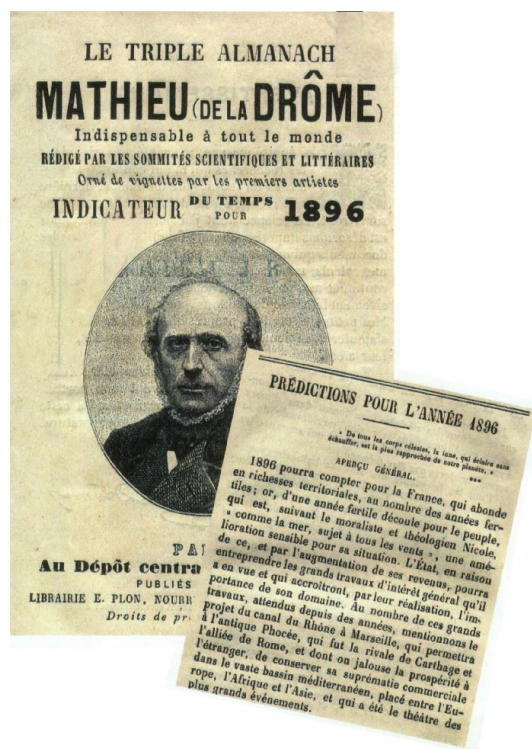
The main goal of this thesis is to study another approach to addressing model uncertainty in the global coupled model CNRM-CM5, by adding stochastic perturbations to the dynamics of the atmospheric model ARPEGE-Climate. Our method, called “stochastic dynamics”, consists in adding perturbations to the temperature, specific humidity and vorticity fields, thus correcting estimations of model initial tendency errors. In this thesis, two initial tendency error estimation techniques were studied, based on nudging the model towards reference data. They yield different results in terms of re-forecast scores, depending on the regions studied.

If the initial tendency error corrections are estimated using an iterative nudging method towards the ERA-Interim reanalysis, seasonal prediction scores over the Northern Hemisphere in winter are significantly improved by drawing random corrections. The 500 hPa geopotential height is also clearly reduced. A re-forecast using “optimal” perturbations drawn within the initial tendency error corrections from the current forecast month shows that useful information at a monthly timescale exists, and could allow significant forecast improvement. The last part of this thesis focuses on the idea of classifying the model perturbations according to its current state during the forecast, in order to take a step closer (if possible) to the improvements noted with these optimal perturbations.

AVANT-PROPOS

L'intérêt des sociétés pour le temps de la saison à venir n'est pas récent. Les méthodes empiriques de prévision saisonnière héritées de traditions ancestrales ne manquent pas, seule leur efficacité réelle laisse à désirer. Si certains cherchent des indices sur la saison à venir dans les épaisseurs de pelures d'oignons, les dictons comme « Noël au balcon, Pâques aux tisons » et les vols d'hirondelles, d'autres basent leurs prévisions sur des méthodes plus élaborées mais parfois tout aussi farfelues. Ainsi, au XIX^{ième} siècle, Philippe-Antoine Mathieu de la Drôme déclarait à l'attention de l'Académie des Sciences :

« On peut prédire le temps comme on prédit le lever et le coucher des astres, dix ans, vingt ans à l'avance »



Inutile de dire que ces propos suscitèrent la colère de l'astronome Urbain Le Verrier, qui avait développé le réseau d'observatoires météorologiques en France. Mathieu de la Drôme en fit toutefois son fond de commerce, et avec succès, puisque ses prévisions climatiques annuelles pour la France furent publiées jusqu'en 1926 sous la forme d'un almanach comme celui représenté ici.¹ Aujourd'hui encore, on trouve sur Internet des prévisionnistes amateurs qui utilisent les taches solaires ou les phases de la Lune pour prévoir le temps qu'il fera pour le défilé du 14 juillet l'an prochain.²

La démarche pseudo-scientifique derrière ces prévisions peut prêter à rire, mais celles-ci soulignent l'intérêt financier et social d'avoir des informations à plus longue échéance sur les conditions météorologiques d'une région. Si ce genre de site existe, c'est avant tout que la faible prévisibilité du climat sur nos latitudes et les faiblesses des modèles de climat ne permettent pour l'heure pas aux

1. Comme le rappelle une réplique d'un film culte : « On peut tromper une fois mille personnes... »
2. J'épargnerai au lecteur les liens vers ces sites d'astrologie moderne, mais la requête *prévision saisonnière* sur un moteur de recherche donne un florilège de sites plus ou moins farfelus, dont un en première position !

centres de prévision nationaux de fournir une information à la fois utile aux décideurs et compréhensible par le grand public.

La prévision saisonnière à l'aide de modèles numériques est toutefois un domaine de recherche reconnu, notamment depuis les succès d'efforts internationaux de prévision sur les régions tropicales. A titre d'exemple, les prévisions de modèles numériques de climat sont utilisées conjointement à des méthodes statistiques sur la région de l'Afrique de l'ouest pour fournir des informations sur la saison de mousson à venir.

Malgré ces réussites, les prévisions des modèles numériques de climat à l'échelle saisonnière souffrent de problèmes récurrents. Ce travail de thèse vise à s'attaquer à une des facettes de la modélisation du climat à l'échelle saisonnière, qui est l'amélioration de la prise en compte des incertitudes des prévisions.

Le manuscrit est organisé en trois parties. La première, à visée introductive, comporte un chapitre de présentation des fondements théoriques de la prévision à cette échéance, et un deuxième chapitre détaillant les outils et méthodes employés au cours de ce travail. La deuxième partie porte sur l'évaluation de l'état de l'art actuel, en prenant l'exemple des prévisions d'ensemble des précipitations sur l'Afrique. La troisième partie de ce manuscrit est séparée en deux chapitres, qui portent sur une méthode appelée « dynamique stochastique » développée au cours de ces trois années de thèse. Enfin, nous terminons ce manuscrit en rappelant les principales conclusions et les perspectives ouvertes par ce travail.

PREMIÈRE PARTIE

INTRODUCTION AUX PRÉVISIONS SAISONNIÈRES NUMÉRIQUES

1

PRÉVISIONS SAISONNIÈRES DYNAMIQUES : FONDEMENTS ET ÉTAT DE L'ART

Ce chapitre introductif établit les bases sur lesquelles repose ce travail de thèse. Il est organisé en trois parties, et débouche sur la formulation des problématiques abordées au cours de cette thèse.

Dans un premier temps, nous abordons les fondements théoriques de la prévision du climat à l'échelle saisonnière, en examinant ce qu'on peut prévoir à cette échelle et pourquoi. Les principaux modes de variabilité du climat à l'échelle saisonnière font l'objet d'une attention particulière. Ensuite, nous voyons les différentes étapes qui ont mené à la modélisation dynamique du climat, telle qu'elle est pratiquée aujourd'hui dans de nombreux centres de recherche. La troisième partie est dédiée à un point au cœur des problématiques de cette thèse : la prise en compte des incertitudes dans les prévisions. Elle permet notamment d'introduire la notion de perturbations stochastiques d'un modèle.

La dernière partie de ce chapitre présente les principales questions auxquelles nous allons essayer de répondre dans ce mémoire, en détaillant l'organisation de la suite du manuscrit, et le lien entre les chapitres et les différentes problématiques étudiées.

1.1 Fondements théoriques de la prévision à longue échéance

Si on parle de « prévisions saisonnières », cela suggère qu'il y a bien quelque chose à prévoir à cette échelle de temps. Le but de cette partie est de présenter les bases théoriques de la prévision saisonnière, en montrant les principales sources de prévisibilité et les liens entre les différentes composantes du climat (océan, atmosphère, glace de mer et continents) qu'on va chercher à représenter à cette échelle. On se focalise ici sur l'atmosphère, qui est au cœur de ce travail de thèse.

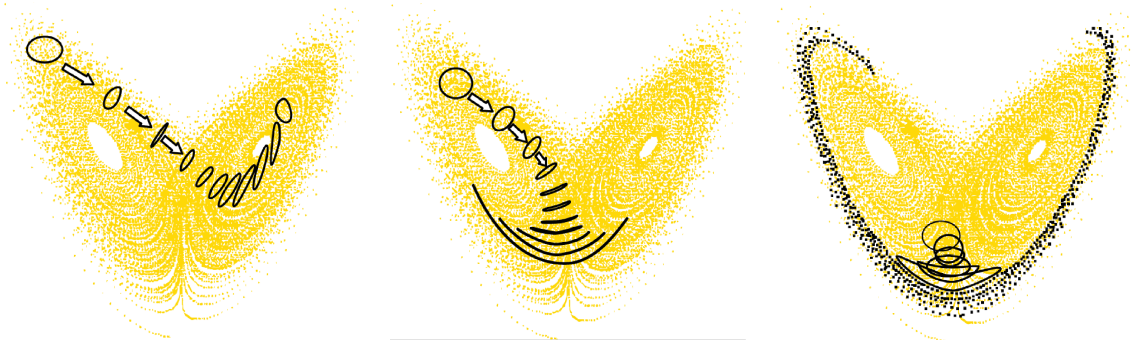


FIGURE 1.1 – Trajectoires dans l'espace des phases solutions du système de Lorenz (1963) : les trois figures montrent l'évolution de points pris dans un voisinage donné autour de trois conditions initiales différentes. (Figures tirées de Palmer, 2009)

1.1.1 Que peut-on prévoir à l'échelle saisonnière ?

1.1.1.1 Théorie du chaos de Lorenz

Une question fondamentale en prévision numérique du temps est celle de la limite de prévisibilité. Les travaux de Edward Lorenz sur ce sujet au cours de la deuxième moitié du vingtième siècle ont été déterminants. En 1963, Lorenz écrit un papier qui marquera la naissance de la théorie du chaos. En étudiant un modèle simplifié de convection atmosphérique, il isole un sous-ensemble de trois équations dont les solutions sont apériodiques et instables (Lorenz, 1963).

$$\begin{cases} \frac{dX}{dt} = \sigma(Y - X) \\ \frac{dY}{dt} = X(r - Z) - Y \\ \frac{dZ}{dt} = XY - bZ \end{cases}$$

En partant d'un état initial très proche de l'origine, Lorenz calcule plusieurs milliers d'itérations du système et étudie les trajectoires obtenues. Il constate que hormis trois solutions stationnaires et un sous-ensemble dénombrable de solutions périodiques, le système admet des solutions qui sont non-périodiques et instables. Ceci se traduit de la manière suivante : deux états proches à une erreur de mesure près divergent rapidement, comme on peut le voir dans la figure 1.1. Les trois cas montrent l'évolution de points pris dans un voisinage restreint avec le modèle de Lorenz. Ceci donne un aperçu du taux d'accroissement des erreurs sur une condition initiale dans un système chaotique. Dans le premier cas, les points restent dans un volume similaire au cours du temps, l'erreur initiale n'a donc que peu d'impact sur l'état final du système. Au contraire, dans le deuxième et *a fortiori* le troisième cas, une erreur sur les conditions initiales rend la détermination de l'état final impossible.

La généralisation de ce constat aux équations d'évolution de l'atmosphère incite Lorenz à formuler l'assertion suivante : soit l'évolution de l'atmosphère est telle qu'elle peut revenir exactement dans le même état qu'auparavant, auquel cas elle obéit à une dynamique quasi-périodique ; soit il n'existe pas d'états parfaitement semblables dans l'histoire de l'évolution de l'atmosphère, et celle-ci a donc une évolution chaotique. Il en découle qu'une simple erreur dans les conditions initiales, quand bien même la modélisation de l'atmosphère serait parfaite,

mène en un temps fini à des différences dans la prévision aussi grandes que si on choisissait un état initial aléatoire. Il existe donc une limite à la prévisibilité de l'atmosphère.

Les travaux de Lorenz ne se limitent pas à ce papier fondateur. Il étudie la propagation d'erreur et le taux de croissance des erreurs dans un modèle de circulation plus réaliste et estime la limite de prévisibilité de l'atmosphère à 17 jours dans ce contexte (Lorenz, 1969). Toutefois, la question de savoir si la théorie du chaos s'applique aux équations de Navier-Stokes reste encore à être résolue sur le plan formel.¹ Mais toutes les études empiriques semblent concorder sur le fait que l'atmosphère obéit à un comportement chaotique.

1.1.1.2 Qu'est-ce qu'une prévision saisonnière ?

Dès lors qu'on admet que l'atmosphère est régie par le chaos, que peut-on espérer prévoir au-delà des limites théoriques de prévisibilité ? Cette limite estimée par Lorenz correspond à un temps caractéristique de croissance d'erreurs sur les conditions initiales. Pour dépasser ce seuil il faut donc s'affranchir de l'impact de ces erreurs initiales, mais avant tout il faut admettre que ce qu'on va prévoir n'est pas l'état de l'atmosphère à un instant t éloigné dans le temps, mais plutôt des statistiques caractérisant l'état de l'atmosphère au cours du mois ou de la saison à venir. On s'intéresse généralement aux températures et aux précipitations moyennes sur trois mois, avec une échéance de un à six mois depuis la date de démarrage des prévisions. Au vu des échelles temporelles en jeu, il est vain de vouloir prévoir température ou précipitations sur un point précis sans utiliser des outils statistiques pointus de descente d'échelle. Voilà pourquoi on regarde généralement ces moyennes sur une zone géographique assez large. Les séries temporelles des différentes variables peuvent apporter des informations sur la distribution de probabilité ou les fréquences de types de temps, mais en aucun cas ne doivent être prises telles quelles. Lorenz nous a prévenus : au-delà de quelques jours un champ atmosphérique à une date donnée dans la prévision ne ressemblera en rien à ce qui aura été réellement observé. Par contre, en regardant les moments statistiques d'une variable donnée, et en les comparant à ceux de la climatologie, on pourra déduire des informations utiles sur la saison à venir.

Les prévisions saisonnières telles que fournies par les différents centres de recherche ou institutions prennent souvent la forme de prévisions probabilistes. Un des héritages des travaux de Lorenz est la prévision d'ensemble, qui consiste à faire un jeu de plusieurs prévisions pour une même saison et un même modèle, afin de prendre en compte les incertitudes liées aux conditions initiales (voir la partie 1.3.1 de ce chapitre). Ces ensembles de prévisions permettent de dresser des probabilités d'occurrence par catégorie de température ou de précipitations (proche, au-dessus ou en dessous de la normale) par simple dénombrement.

La figure 1.2 montre les prévisions saisonnières de précipitation pour la saison juin-juillet-août (JJA) à un mois d'échéance émises en juin 2012 par Météo-France. Le code couleur indique pour chaque point de grille la catégorie de précipitations la plus probable, et l'intensité des couleurs la probabilité obtenue pour cette catégorie. A titre d'exemple, l'ensemble prévoit des précipitations inférieures à la normale avec une forte probabilité au-dessus de la Guyane. Les zones blanches sont celles pour lesquelles on ne peut pas favoriser une catégorie par rapport à une autre. La lecture de ce genre de carte n'est possible qu'au travers du prisme des

1. C'est d'ailleurs un des aspects de la compréhension des solutions des équations de Navier-Stokes qui constitue l'une des questions fondamentales des mathématiques d'après l'Institut Clay.

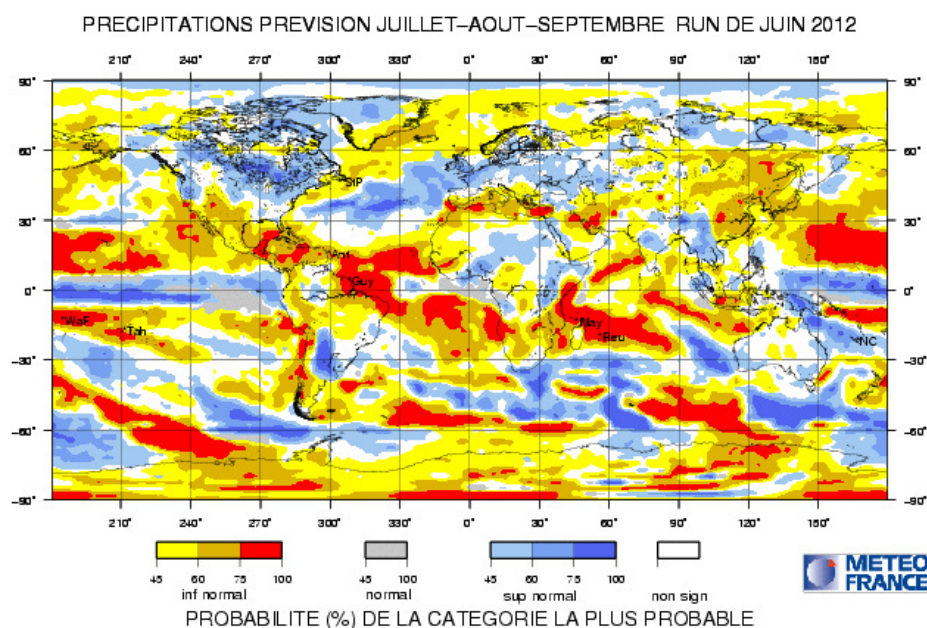


FIGURE 1.2 – Exemple de carte de synthèse des prévisions probabilistes saisonnières pour les précipitations en juin-juillet-août 2012 (prévision de juin 2012) par Météo-France. (Source : Direction de la Climatologie, Météo-France)

performances passées du modèle. Par exemple, les prévisions saisonnières de précipitation sont très peu performantes sur la plupart des régions extra-tropicales, et les indications données sur la figure 1.2 sont donc à prendre avec précaution sur ces régions.

Si l'atmosphère est un système chaotique, certaines régions s'avèrent peu sensibles aux fluctuations de petite échelle. Sur les régions tropicales, en particulier au-dessus des océans, la circulation de grande échelle et les précipitations dépendent en grande partie des conditions océaniques et en premier lieu des températures de surface de l'océan (TSO). Des expériences de prévision saisonnière utilisant des modèles atmosphériques globaux avec différents jeux de conditions initiales atmosphériques et de forçages océaniques ont montré que sur les tropiques, l'océan jouait un rôle primordial. Deux simulations avec des conditions initiales atmosphériques très différentes convergent rapidement vers des états similaires sur les tropiques si les conditions océaniques sont les mêmes (Shukla, 1998). Pour les régions tempérées et les hautes latitudes, le constat est différent. On peut espérer s'affranchir du chaos les années où les signaux océaniques sont très marqués (Trenberth et al., 1998), mais la prévisibilité de l'atmosphère demeure plus faible sur ces régions.

Pour faire des prévisions à l'échelle saisonnière, les modèles doivent donc prendre en compte les interactions de grande échelle spatio-temporelle entre l'atmosphère et d'autres composantes du système Terre qui sont moins chaotiques, comme les surfaces continentales ou l'océan. Ils doivent aussi représenter les sources de prévisibilité internes comme la stratosphère.

1.1.2 Modes de variabilité à l'échelle saisonnière

L'intérêt des scientifiques pour la prévision à l'échelle saisonnière est sans doute antérieur au développement des prévisions à court terme, accéléré par l'avènement des calculateurs

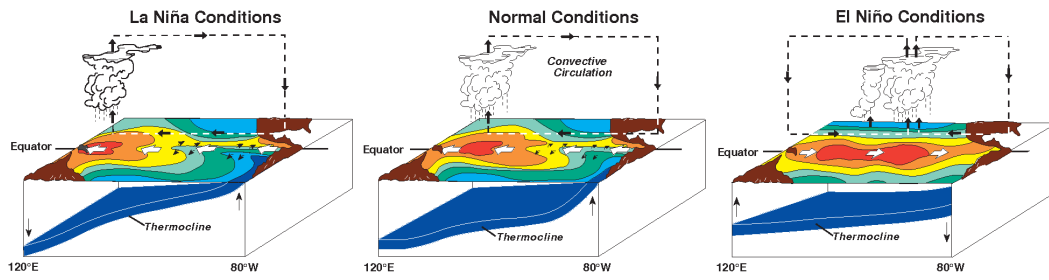


FIGURE 1.3 – Position de la thermocline et des zones de convergence, et sens dominant des courants équatoriaux pendant les phases « La Niña », normales et « El Niño » de l'ENSO. (Source : NOAA/PMEL/TAO)

numériques (Troccoli, 2010). Le développement de la prévision saisonnière a découlé de l'étude et la compréhension des modes de variabilité du climat aux échelles globale et saisonnière, et des liens appelés téléconnexions entre phénomènes climatiques se développant à des endroits distants sur le globe. Les principales téléconnexions auxquelles nous nous sommes intéressés pendant cette thèse sont détaillées ci-dessous.

1.1.2.1 L'Oscillation Australe - El Niño (ENSO)

Dès le début du vingtième siècle, Sir Gilbert Walker découvrit un lien, qu'il appela Oscillation Australe (Southern Oscillation), entre les pressions atmosphériques sur le Pacifique Sud et l'Indonésie. Walker était alors Directeur Général des Observatoires en Inde, et une de ses missions était de prévoir la mousson indienne afin de pouvoir anticiper les conséquences potentiellement catastrophiques des fluctuations de mousson. En cherchant des corrélations du champ de pression sur les régions intertropicales, il mit en évidence cette oscillation sans toutefois parvenir à l'expliquer. Ses travaux, publiés en 1924, sont souvent considérés comme le fondement de la prévision saisonnière.

Utilisée comme prédicteur statistique de la mousson indienne, l'Oscillation Australe n'était pas le « Graal » espéré et les recherches de Walker furent oubliées pendant quelques décennies avant que les travaux de Bjerknes en 1969 ne fassent le lien entre l'Oscillation Australe et le phénomène appelé « El Niño », caractérisé par un réchauffement interannuel des températures de surface de l'océan (TSO) le long des côtes péruviennes et équatoriennes. Le nom El Niño (l'enfant en espagnol) fait référence à l'apparition de ce réchauffement habituellement aux alentours de Noël (Trenberth, 1997). Ces deux aspects d'un même phénomène décrit ci-dessous ont donné lieu à l'acronyme ENSO (El Niño Southern Oscillation) (Neelin et al., 1998). L'importance du couplage entre l'océan et l'atmosphère dans le Pacifique tropical, mais également les impacts des anomalies de TSO sur le climat à l'échelle saisonnière étaient alors mis en évidence. L'ampleur des événements El Niño de 1982-83 et 1997-98 engendra un grand intérêt scientifique pour le phénomène et ses conséquences, mais associa également le terme El Niño à la variabilité du climat dans l'esprit du grand public.

L'ENSO est un phénomène irrégulier, les différents épisodes d'El Niño survenant à entre 3 et 7 ans d'intervalle, et durant de quelques mois à plus d'un an. La figure 1.3 schématise

les conditions atmosphériques et océaniques sur le Pacifique tropical pendant les deux phases marquées de l'ENSO, El Niño et son pendant La Niña, ainsi que les conditions normales. En conditions normales, la circulation de Walker est alimentée par un contraste de TSO entre l'est (plus froid) et l'ouest du bassin Pacifique tropical, qui engendre un courant d'est à la surface de l'océan. La zone de convergence inter-tropicale (ZCIT) se situe à l'ouest du bassin au niveau de la branche ascendante de la cellule de Walker et est à l'origine de précipitations sur la région. Les événements El Niño sont caractérisés par un affaiblissement des alizés (vents d'est proches de la surface qui constituent la branche basse de la circulation de Walker sur la région), ce qui inverse le sens des courants de surface et induit un réchauffement des TSO sur l'est du bassin. La cellule de Walker est perturbée, et la ZCIT est déplacée vers le centre du Pacifique. Le long des côtes d'Amérique du Sud, le réchauffement de la TSO est accompagné d'un affaiblissement des remontées d'eaux froides (upwelling) et donc d'une chute de la thermocline.

Au contraire, lorsque la cellule de Walker s'intensifie, les courants sont également renforcés et poussent la ZCIT vers l'extrémité ouest du bassin. A l'est, la thermocline remonte et le phénomène d'upwelling est renforcé : ce sont les conditions La Niña.

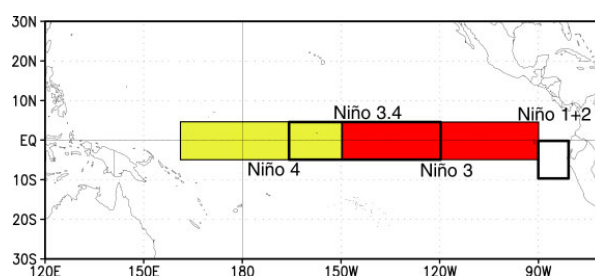


FIGURE 1.4 – Régions utilisées pour la définition des indices Niño. (Source : NOAA/PMEL/TAO)

Plusieurs indices statistiques sont couramment utilisés pour déterminer les phases de l'ENSO. Le plus ancien, qui renvoie aux travaux de Walker, est le Southern Oscillation Index (SOI). Il mesure les différences de pression de surface entre Darwin et Tahiti. Les séries de pression sont généralement lissées et moyennées à l'échelle mensuelle afin d'éviter de prendre en compte d'autres phénomènes à l'échelle intra-mensuelle comme l'oscillation de Madden-Julian (Trenberth, 1997). Grâce au développement des observations globales, d'autres indices peuvent être utilisés, comme par exemple l'anomalie de température de surface sur des régions prédéfinies montrées figure 1.4. Nous utiliserons dans cette thèse les températures sur la région Niño 3.4 comme indicateur de la phase de l'ENSO.

Les impacts de l'ENSO sur le climat sont globaux et durables du fait de la plus grande inertie thermique de l'océan, qui s'ajuste plus lentement à la variabilité de l'atmosphère. Les perturbations engendrées dans l'atmosphère se répercutent à l'échelle globale *via* la propagation d'ondes atmosphériques de grande échelle. Il serait vain de lister ici l'ensemble des impacts couramment observés lors des phases de l'ENSO, mais la figure 1.5 en illustre les principaux en termes de précipitations et de température de surface. Les effets de l'ENSO sont ressentis des latitudes tropicales aux latitudes extra-tropicales, en été comme en hiver.

La part conséquente de l'ENSO dans la variabilité du climat explique les nombreux efforts de recherche portant sur ce phénomène au cours des dernières décennies. Ces efforts ont été appuyés par la multiplication des techniques d'observation de l'océan et de l'atmosphère, avec l'arrivée de la télé-détection par satellite et la mise en œuvre de réseaux de mesures dans l'océan Pacifique. Citons notamment le réseau de bouées ancrées TAO/TRITON (Tropical At-

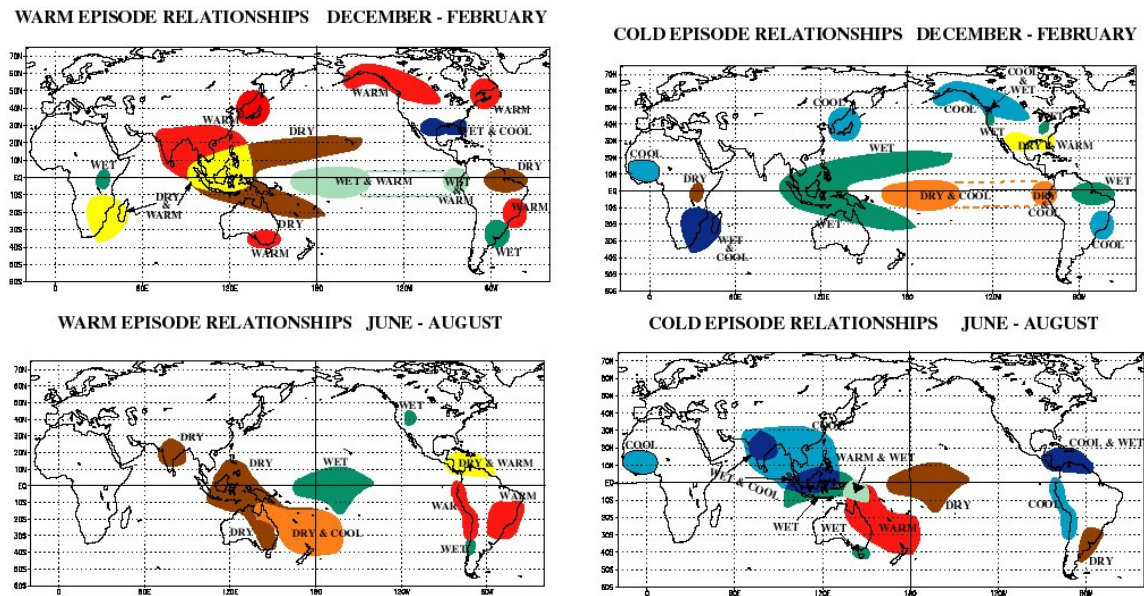


FIGURE 1.5 – Principaux impacts des phases chaudes (à gauche) et froides (à droite) de l'ENSO sur les conditions climatiques des différentes régions du globe, en hiver (en haut) et en été boréal (en bas). (Source : NOAA)

mosphere and Ocean, initié en 1999) qui fournit des données océanographiques pour environ 70 mouillages dans l'océan Pacifique tropical, et le satellite TOPEX/POSEIDON et ses successeurs JASON-1 et JASON-2 qui fournissent des données d'altimétrie de l'océan, de hauteur de vagues et de vitesse des vents sur l'ensemble du globe depuis 1992. Cet essor des techniques d'observation a permis de mieux comprendre le phénomène, même si aujourd'hui plusieurs théories sont avancées pour caractériser cette oscillation et expliquer son irrégularité. Hoskins et Schopf (2008) les classent en trois catégories : la théorie du phénomène auto-entretenu, celle de l'oscillateur amorti, où l'influence du forçage aléatoire de l'atmosphère prédomine, et une théorie hybride où l'ENSO est un phénomène aux caractéristiques temporelles provenant d'un mode dominant oscillatoire, mais où l'influence du bruit est telle qu'elle perturbe la régularité des oscillations.

Malgré l'importance du phénomène, la prévision saisonnière ne peut reposer entièrement sur une prévision correcte de l'ENSO. En témoignent les effets habituels de El Niño attendus en 1997–1998 qui n'eurent pas lieu (sécheresses annoncées sur l'Inde ou en Afrique du sud, peut-être atténuées par des anomalies de TSO simultanées sur l'océan Indien), ce qui nuit à l'aura de la prévision saisonnière dans l'opinion publique des régions concernées. Ces échecs de la prévision saisonnière uniquement basée sur El Niño montrent que d'autres sources de variabilité climatique doivent être correctement représentées dans les systèmes de prévision.

1.1.2.2 L'Oscillation Nord-Atlantique (NAO)

L'Oscillation Nord-Atlantique, ou NAO, est le premier mode de variabilité de la troposphère sur l'Atlantique Nord extra-tropical, sur une vaste plage allant de la décennie à la décennie. Elle se traduit par des fluctuations de l'intensité d'un dipôle de pression entre l'anticyclone des Açores et la dépression islandaise (Walker et Bliss, 1932). Ces modifications du gradient

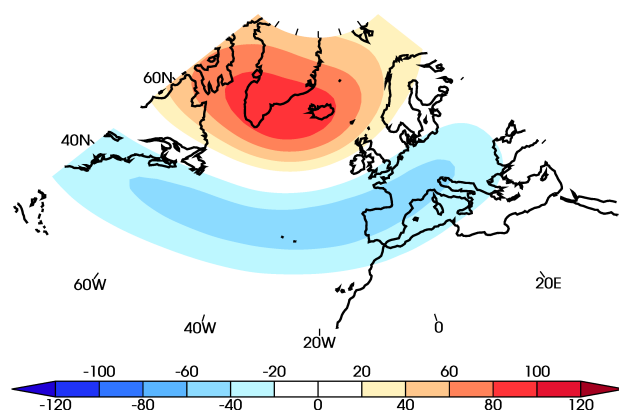


FIGURE 1.6 – Premier mode de variabilité des anomalies mensuelles de géopotential à 500 hPa sur la zone Euro-Atlantique : le pattern NAO pour l'hiver boréal. Données ERA-Interim pour Décembre-Janvier-Février (DJF) 1979 à 2010.

de pression modulent l'intensité et la direction du courant jet dirigé vers l'Europe, le rail des dépressions et la répartition des précipitations sur le continent européen. En phase positive de la NAO (NAO⁺), l'anticyclone des Açores et la dépression islandaise sont accentués, le gradient de pression est donc plus élevé et la position du courant jet et du rail de dépressions induit des conditions plus humides et douces sur le nord de l'Europe. La Méditerranée connaît à l'inverse des conditions plus sèches. Sur le continent américain, l'accentuation de la dépression islandaise provoque l'arrivée d'air froid et sec sur le Québec et le Groenland, tandis que les hivers sont plus doux et humides aux Etats-Unis. En phase NAO⁻, le courant jet est plus méridional du fait de l'affaiblissement des deux centres d'action, et les dépressions sont donc décalées vers le sud. Il en résulte une inversion des conséquences en termes d'humidité et de température sur l'Europe et les Etats-Unis.

Le mode NAO est une part importante de la variabilité mensuelle tout au long de l'année, toutefois ce mode est particulièrement marqué en hiver. A l'échelle mensuelle, on peut définir un indice NAO en projetant l'anomalie mensuelle de géopotential sur le « pattern » montré à la figure 1.6. Nous reviendrons sur la définition de cet indice dans le chapitre suivant.

L'hiver 2009-2010 a marqué les esprits par ses conditions météorologiques froides sur l'Europe et plusieurs épisodes neigeux intenses. Ceci était lié à la plus longue persistance de conditions NAO⁻ jamais enregistrée depuis l'hiver 1939-1940 (Ouzeau et al., 2011). Prévoir correctement un indice NAO à l'échelle saisonnière pourrait permettre d'anticiper les conséquences néfastes d'épisodes neigeux sur l'Europe de l'ouest, mais il y a encore beaucoup de chemin à faire. La prévision de l'indice NAO plutôt qu'un champ de géopotential permet d'obtenir des scores significativement positifs mais qui demeurent néanmoins assez faibles (Doblas-Reyes et al., 2003).

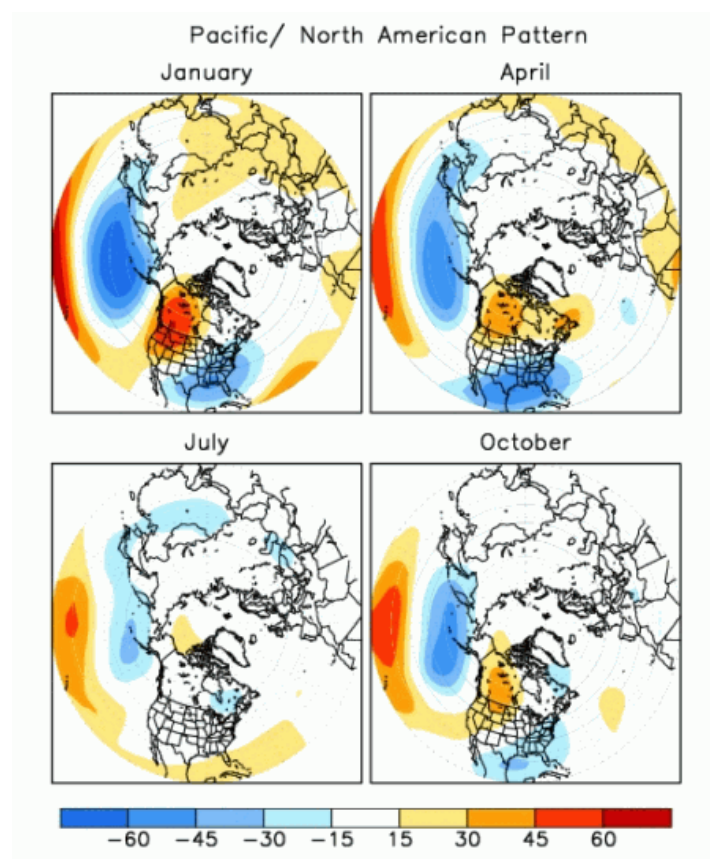


FIGURE 1.7 – Pattern de la phase positive du PNA pour différents mois de l’année (corrélation temporelle entre anomalies de géopotential et l’indice PNA du CPC). (Source : NOAA/NCEP)

1.1.2.3 La téléconnexion Pacifique Amérique du Nord (PNA)

Si la variabilité du climat sur l’Atlantique Nord à l’échelle mensuelle est dominée par un pattern zonal, la NAO, sur la région de l’Amérique du Nord et de l’océan Pacifique Nord, le premier mode de variabilité prend la forme d’un train d’ondes en arc de cercle des tropiques vers les moyennes latitudes. Ce signal, appelé PNA (pour « Pacific/North American pattern »), est particulièrement marqué dans la moyenne troposphère par des anomalies d’altitude de géopotential (Wallace et Gutzler, 1981).

Lors de la phase positive du mode PNA, on observe des anomalies positives d’altitude de géopotential sur les régions proches de Hawaii et des Grandes Plaines nord-américaines, et des anomalies négatives au sud des îles Aléoutiennes et sur le Sud-Est des Etats-Unis, liées à une amplification du courant-jet est-asiatique. La figure 1.7 montre le pattern de la phase positive de PNA pour différents mois de l’année, sous la forme d’une corrélation temporelle à chaque point de grille entre les anomalies standard de géopotential et la série de l’indice PNA du Climate Prediction Center.

L’influence de l’ENSO sur le mode PNA a été mise en évidence dans les observations (Horel et Wallace, 1981) et repose sur la propagation du Pacifique tropical vers les moyennes latitudes d’ondes de Rossby (Trenberth et al., 1998), incurvées vers l’est du fait de la force de Coriolis. Les épisodes El Niño sont ainsi souvent accompagnés de phases positives du mode PNA, tandis

qu'au cours d'épisodes La Niña, c'est au contraire la phase négative qui est favorisée.

1.1.2.4 Les modes annulaires des hémisphères Nord et Sud (NAM et SAM)

L'oscillation Nord-Atlantique est parfois vue comme l'empreinte à l'échelle régionale d'un phénomène à l'échelle hémisphérique. Ce phénomène, appelé Oscillation Arctique (AO), est caractérisé par des fluctuations de l'intensité du vortex polaire hivernal aux échelles mensuelle et interannuelle, et se traduit en surface par un phénomène de balancement de la pression atmosphérique entre les régions de l'Arctique et les latitudes tempérées. La signature de cette oscillation dans la basse stratosphère comporte une symétrie zonale prépondérante, et on observe le même phénomène dans l'hémisphère Sud. Aussi Thompson et Wallace (2000) ont-ils proposé de nommer ces phénomènes « Northern Annular Mode » et « Southern Annular Mode », respectivement notés NAM et SAM par la suite.

L'impact du NAM sur le temps est parfois contesté. Nous n'entrerons pas ici dans ce débat concernant les liens entre NAO et AO et leur importance relative, qui a donné lieu à de nombreuses publications au cours de la dernière décennie (par exemple Ambaum et al., 2001; Deser, 2000) mais dépasse très largement le cadre de cette thèse.

1.1.2.5 Sources de prévisibilité internes de l'atmosphère

On ne peut évoquer la variabilité du climat à l'échelle saisonnière à interannuelle sans parler des sources de prévisibilité internes de l'atmosphère, dont certaines sont encore souvent paramétrisées dans les modèles de climat.

Les travaux de Baldwin et al. (2003) ont montré que l'influence de la stratosphère sur la dynamique de la troposphère était déterminante à l'échelle mensuelle. En particulier, une signature de l'ENSO dans la stratosphère a été mise en évidence par van Loon et Labitzke (1987). Une autre source de prévisibilité est l'oscillation quasi-biennale (QBO). Ce mode de variabilité des vents zonaux équatoriaux dans la stratosphère a une fréquence d'environ 28 mois et influe sur la circulation troposphérique de l'hémisphère Nord en hiver pendant au moins un mois et les phénomènes de mousson (Baldwin et al., 2001).

Une autre source de prévisibilité de l'atmosphère est la distribution des aérosols atmosphériques, qui a un impact non négligeable sur le bilan radiatif et peut influencer la circulation atmosphérique aux échelles globales (dans le cas des éruptions volcaniques majeures) ou plus régionales (comme par exemple les poussières sahéliennes, voir Engelstaedter et Washington (2007)).

1.1.3 Autres composantes du système climatique à l'échelle saisonnière

Les téléconnexions évoquées dans la partie précédente sont caractérisées par des variables atmosphériques et océaniques. Toutefois, l'influence d'autres composantes du système climatique global n'est pas à négliger. Les prochains paragraphes visent à donner un aperçu des sources de prévisibilité aux échelles saisonnières à interannuelles contenues dans les interactions entre surfaces continentales et cryosphère, et le système climatique global.

Les surfaces continentales jouent un rôle important dans la mémoire du climat aux échelles intra-saisonnières à interannuelles, par leur influence sur l'équilibre énergétique de l'atmosphère. La composition du sol (type de végétation, couverture neigeuse) modifie son albédo (et donc la part du flux solaire incident réfléchi vers l'atmosphère) ainsi que les flux de chaleur sensible et latente. L'humidité du sol est également une donnée essentielle du cycle hydrologique, et peut influencer la température de surface et les précipitations à une échelle régionale.

1.1.3.1 Le rôle de l'humidité des sols

Les anomalies d'humidité du sol pouvant persister jusqu'à plusieurs mois, l'état du sol est au même titre que l'océan une composante lente du système climatique pouvant servir de « mémoire » à l'échelle de temps saisonnière. Dans le cadre du projet GLACE (Global Land-Atmosphere Coupling Experiment), Koster et al. (2004) ont montré à l'aide de modèles numériques que l'humidité du sol avait une forte influence sur les précipitations en été dans les régions intermédiaires entre des climats humides et des climats secs, là où l'évaporation est assez sensible aux anomalies d'humidité du sol et où elle est assez élevée pour provoquer des précipitations. Les résultats de Koster et al. (2004) ont généralisé ceux obtenus par plusieurs études avec des modèles indépendants (voir Dirmeyer, 1999; Douville et Chauvin, 2000). Une conclusion fondamentale de Koster et al. (2004) était qu'il est important d'inclure une bonne modélisation de l'humidité des sols dans les modèles utilisés en prévision saisonnière, afin d'améliorer les prévisions des précipitations.

Un deuxième jeu d'expériences dans le cadre de GLACE2 (Koster et al., 2010) s'est focalisé sur des échéances intra-saisonnières et a permis d'affiner ces résultats, en évaluant la prévisibilité potentielle liée à une bonne initialisation de l'humidité des sols dans des modèles de climat. L'apport potentiel de la prise en compte des conditions initiales d'humidité des sols pour les prévisions de température de surface et des précipitations est plus marqué sur des régions comme l'Amérique du Nord, soumises à un climat plus continental où les processus de surface prennent une plus grande importance, qu'en Europe, où les anomalies sont plutôt régies par la circulation atmosphérique de grande échelle (van den Hurk et al., 2012). D'autres études montrent malgré tout une prévisibilité potentielle liée à l'humidité des sols en Europe (Douville, 2010).

1.1.3.2 L'influence de la cryosphère

La cryosphère désigne l'ensemble des régions du globe où l'eau est présente à l'état solide : banquise, glaciers, calottes polaires, mais également les sols recouverts de neige ou gelés (de manière permanente ou non). Elle représente, selon la saison, de 7 à 17% de la surface de la Terre.

La neige a des propriétés thermodynamiques qui en font un acteur majeur du climat aux hautes latitudes. L'albédo de la neige est élevé, la part de rayonnement solaire réfléchi est donc grande. De plus, la neige a une émissivité proche d'un corps noir, et une forte chaleur spécifique. Il en résulte une rétro-action positive sur le climat local, car la présence de neige favorise le refroidissement des températures de surface. De nombreuses études ont mis en évidence l'influence de la neige sur le climat, *via* notamment des téléconnexions atmosphériques. Pour ne citer que deux exemples, Peings et Douville (2009) ont étudié l'influence de la couverture de neige du continent eurasiatique sur la mousson indienne l'été suivant, et Cohen

et Entekhabi (1999) ont proposé un mécanisme liant la couverture de neige en automne en Sibérie et l'Oscillation Arctique.

Un autre aspect de l'influence de la cryosphère sur le climat à l'échelle interannuelle est le rôle de la glace de mer dans les échanges énergétiques entre l'atmosphère et l'océan. Là encore, la présence de glace de mer modifie profondément la nature des flux d'énergie à la surface de la Terre. Le réchauffement global du climat au cours des dernières décennies a incité des chercheurs à s'intéresser à l'influence de la présence ou l'absence de glace de mer sur la circulation atmosphérique. Royer et al. (1990) ont étudié la sensibilité d'un modèle d'atmosphère à la présence de glace de mer, en simulant l'effet d'un océan libre de glace aux frontières du modèle, par rapport à une simulation de contrôle incluant le forçage de la glace de mer.

Plus récemment, l'accélération de la réduction de la superficie de glace de mer en Arctique au minimum de septembre a donné lieu à de nombreuses études montrant l'influence d'anomalies de couverture de glace de mer en été sur les vagues de froid en Asie (Honda et al., 2009) ou plus généralement sur la circulation atmosphérique (Francis et al., 2009) l'hiver suivant. Chevallier et Salas y Mélia (2012) ont montré une prévisibilité potentielle de ces anomalies de couverture de glace en été jusqu'à 6 mois à l'avance. Ces études soulignent l'importance de l'inclusion d'une composante de glace de mer dans les modèles de climat (Orsolini et al., 2012).

1.1.3.3 Limites pratiques à la prévision de ces composantes

Malgré les nombreuses études mettant en évidence des liens forts entre ces composantes du système climatique et l'évolution des variables atmosphériques à l'échelle saisonnière, un défi majeur avant d'inclure de manière efficace ces données dans un modèle est celui de l'amélioration des jeux d'observations. L'initialisation des modèles avec des données fiables est particulièrement délicat pour l'humidité des sols, la neige et la glace de mer évoquées dans ce paragraphe.

Synthèse

Cette première partie nous a permis de présenter ce qu'on entend par prévision saisonnière, et les principales sources de prévisibilité du climat à cette échelle.

La théorie du chaos introduite par Lorenz (1963) démontre qu'il existe une limite de prévisibilité de l'atmosphère (au sens où on l'entend généralement pour les prévisions météorologiques) de l'ordre de la dizaine de jours. Les prévisions saisonnières se situent au-delà de cette limite, et cherchent donc généralement à informer l'utilisateur sur des anomalies moyennes sur une saison à venir des températures et/ou précipitations. Ces prévisions sont plus performantes sur les régions tropicales, mais certaines sources de prévisibilité permettent d'envisager des prévisions sur les régions tempérées.

Le principal mode de variabilité du climat à l'échelle saisonnière est l'ENSO, dont les effets se font sentir à l'échelle globale. Il existe d'autres modes de variabilité plus régionaux, comme l'oscillation Nord-Atlantique ou la téléconnexion Pacifique Amérique du Nord. Ces phénomènes ont une signature à la fois atmosphérique et océanique. Il existe des sources de prévisibilité internes à l'atmosphère, comme l'influence de la stratosphère. Les autres composantes du

système climatique (surfaces continentales, glace de mer) jouent également un rôle important dans la variabilité du climat à l'échelle saisonnière.

La compréhension de ces différents modes de variabilité et sources de prévisibilité a précédé puis accompagné le développement de modèles numériques de climat, dont un aperçu historique est dessiné dans la partie suivante.

1.2 De la compréhension de l'ENSO à la modélisation numérique du climat

Dans cette partie, nous présentons les grandes étapes du développement des modèles numériques du climat, maintenant fréquemment utilisés pour la prévision saisonnière. Nous revenons pour cela sur l'émergence des modèles et le développement de nouvelles techniques de modélisation, en évoquant notamment des projets de recherche internationaux qui ont permis des avancées notables dans le domaine.

1.2.1 Prévisions statistiques à l'échelle saisonnière

Les liens détaillés dans la partie précédente sont souvent utilisés dans le but de prévoir à l'échelle saisonnière certains phénomènes climatiques à l'aide de méthodes statistiques. De nombreux centres de prévision du climat utilisent des méthodes statistiques pour compléter (ou remplacer) l'utilisation de modèles numériques plus complexes. De nombreuses études soulignent la compétitivité de ces méthodes statistiques vis-à-vis des modèles numériques, par exemple pour prévoir l'ENSO (Barnston et al., 1999; van Oldenborgh et al., 2005), les phénomènes de mousson (Fontaine et al., 1999; Mo et Thiaw, 2002; Rajeevan et al., 2007) ou encore le climat des latitudes tempérées (Cohen et Fletcher, 2007) à l'échelle saisonnière. Pour une revue détaillée des techniques de prévisions saisonnières statistiques, on pourra se référer à Mason et Baddour (2008).

Toutefois, chaque modèle statistique est développé et optimisé pour un phénomène climatique particulier à une échelle donnée. L'inconvénient de ces méthodes est leur très forte dépendance de la qualité des données d'observation utilisées lors de la période d'apprentissage, et leur vulnérabilité à l'heure d'une évolution rapide du climat.² Les liens statistiques établis sur une période des 30 ou 40 dernières années entre plusieurs variables physiques du climat ne sont pas nécessairement robustes pour expliquer les liens pour la saison à venir, et certaines études montrent qu'il vaut mieux parfois construire un modèle sur les 10 dernières années observées plutôt que sur l'ensemble de la période de données dont on dispose (Huang et al., 1996; Livezey et al., 2007).

2. Comme l'a rappelé Michel Jarraud, secrétaire général de l'OMM, en clôture du séminaire annuel du CEPMMT portant cette année sur la prévision saisonnière : « The past is no longer a good indicator for the future ».

1.2.2 L'apparition des modèles numériques de climat

En parallèle de la compréhension des phénomènes physiques de grande échelle spatio-temporelle, l'intégration de modèles de circulation générale (MCG) à partir de conditions initiales réalistes se développa au cours des années 1970-1980, avec la réalisation des premières analyses globales opérationnelles aux Etats-Unis et en Europe, et les premières simulations dites « à longue échéance ».

L'apparition de modèles numériques de climat a sans doute été retardée par la relégation au second plan des études statistiques initiées par Sir Gilbert Walker et ses contemporains, au moment de l'émergence de modèles numériques de prévision du temps à courte échéance. Les travaux de Richardson dans les années 1920, puis de Rossby, Eliassen, Charney, Bjerknes et von Neuman dans les années 1940 et 1950 (pour ne citer que quelques grands noms), ont contribué au développement des modèles numériques de prévision du temps utilisés aujourd'hui. La théorie de Lorenz a toutefois conforté les chercheurs dans l'idée qu'utiliser des modèles numériques de prévision pour des échéances allant au-delà de quelques jours serait parfaitement inutile.

Au même moment que les travaux fondateurs de Lorenz, Manabe et al. (1965) montrent qu'il est possible d'intégrer un modèle de circulation générale atmosphérique pendant une période de plusieurs mois, et que le climat modélisé comporte certaines caractéristiques cohérentes avec les données d'observation. Mais ce sont les travaux de Shukla (1981) qui sont très souvent considérés comme étant à l'origine de la prévision saisonnière avec les MCG. Dans cet article, J. Shukla intègre un modèle de circulation générale atmosphérique neuf fois pendant 60 jours en partant de neuf conditions initiales atmosphériques correspondant à trois 1^{er} janvier successifs avec des perturbations initiales. Le modèle est forcé aux frontières par des températures de surface de l'océan, de la glace de mer, et des conditions de surfaces continentales climatologiques et identiques pour chaque intégration. En étudiant les moyennes dans le temps et l'espace des variables modélisées, J. Shukla se concentre sur la prévisibilité due aux forçages externes. Le choix de conditions initiales séparées d'un an lui permet d'évaluer au bout de combien de temps la condition initiale atmosphérique est oubliée, de la même manière que l'étude de Lorenz (1969) à l'échelle synoptique. Les résultats montrent qu'il existe une prévisibilité de l'atmosphère aux échéances de 2 semaines à un mois, et que au-delà, la prévisibilité disparaît du fait de trop grandes erreurs du modèle. Ces conclusions ouvrent la voie à de nombreux travaux de recherche en prévision numérique à plus longue échéance.

1.2.3 La prévision saisonnière numérique

1.2.3.1 Principe et développement de la prévision numérique

On a vu précédemment que les composantes lentes qu'étaient l'océan, la surface des continents et la glace de mer expliquaient une part primordiale de la variabilité du climat à l'échelle saisonnière. Très tôt, les chercheurs en modélisation ont donc essayé de représenter non seulement l'atmosphère mais ces différentes composantes, avec des degrés de complexité différents. La plupart des travaux de recherche à la fin du XX^{ième} siècle se sont focalisés sur la prévision de l'ENSO (Goddard et al., 2001). C'est Zebiak et Cane (1987) qui ont réussi en premier lieu à représenter l'ENSO avec un modèle couplé océan-atmosphère simplifié, de manière à prévoir l'épisode El Niño de 1986-1987 avec presque un an d'avance.

Aujourd'hui, les modèles de circulation générale utilisés en prévision saisonnière de l'atmosphère entrent le plus souvent dans deux catégories :

1. les modèles à deux étapes, ou « **two-tiered** » (Bengtsson et al., 1993), pour lesquels on prévoit d'abord les conditions aux bords (océan, surfaces continentales, glace de mer) à l'aide de méthodes statistiques ou dynamiques, puis on force un modèle global d'atmosphère avec ces conditions ;
2. les modèles à une seule étape, ou « **one-tiered** » où toutes les composantes sont couplées et évoluent simultanément : on parlera de modèles de circulation générale couplés.

L'avantage évident des modèles couplés est d'incorporer l'ensemble des interactions entre l'atmosphère et les autres composantes du système climatique. L'utilisation de systèmes de prévision two-tiered suppose implicitement de négliger la rétroaction de l'atmosphère sur l'évolution des composantes plus lentes du climat. Un modèle couplé pourra donc en théorie reproduire l'évolution du climat à l'échelle saisonnière de manière plus réaliste. Toutefois, ces modèles couplés sont plus coûteux en ressources de calcul, tant pour l'intégration du modèle qu'au cours de la phase d'assimilation de données atmosphériques et océaniques, et certains centres opérationnels préfèrent encore utiliser des modèles atmosphériques forcés pour leurs prévisions à l'échelle saisonnière (Mason, 2008).

1.2.3.2 Les projets de recherche internationaux

Le développement des techniques de modélisation numérique du climat à l'échelle saisonnière à interannuelle telles qu'utilisées aujourd'hui a fait l'objet de plusieurs projets internationaux qui ont permis d'évaluer méthodes et modèles de prévision.

A la fin des années 1990, les projets PROVOST (PRediction Of climate Variations On Seasonal to interannual Time-scales) en Europe et DSP (Dynamical Seasonal Prediction) aux Etats-Unis ont coordonné des expériences de rétro-prévisions saisonnières avec plusieurs modèles globaux atmosphériques forcés avec des TSO « parfaites » (Palmer et Shukla, 2000). Ces projets offrent un cadre d'étude commun permettant d'étudier les performances relatives des différents modèles ainsi que celles d'une combinaison des prévisions de plusieurs modèles (ou prévision multi-modèle, voir le paragraphe 1.3.2.1). Ils permettent d'estimer la part de la prévisibilité due à la composante océanique par rapport aux conditions initiales (Brankovic et Palmer, 2000), mais également de mettre en évidence l'intérêt de l'approche multi-modèle en prévision saisonnière (Doblas-Reyes et al., 2000).

Les conclusions de ces projets incitent à poursuivre de telles études, mais avec des modèles couplés océan-atmosphère. Ce sera notamment le cadre du projet européen DEMETER (Palmer et al., 2004), pour lequel sept modèles couplés océan-atmosphère européens (dont celui du CNRM) sont intégrés sur une période commune de 22 ans, en initialisant neuf prévisions de six mois par modèle, en février, mai, août et novembre de la période. Outre l'utilisation de modèles couplés, la nouveauté du projet DEMETER réside en la volonté d'évaluer les prévisions saisonnières d'un point de vue plus adapté à l'utilisateur final, en utilisant par exemple des techniques de descente d'échelle, des modèles de rendements agricoles ou de risques sanitaires. Le projet DEMETER a confirmé l'intérêt de l'approche multi-modèle (Hagedorn et al., 2005) et a incité plusieurs centres de prévision européens (le MetOffice anglais, le CEPMMT et Météo-France) à produire un ensemble de prévisions saisonnières en temps réel dans le cadre d'un consortium baptisé EUROSIP (EUROpean Seasonal-to-Interannual Prediction).

Le projet européen ENSEMBLES (Doblas-Reyes et al., 2009) a permis d'étudier et de comparer différentes méthodes de prise en compte des incertitudes dans les dernières générations des modèles couplés en prévision saisonnière. On reviendra plus en détail sur ce point dans la partie 1.3.

1.2.4 Techniques de correction des modèles

La prévision à l'échelle saisonnière à l'aide de modèles couplés de circulation générale ne se limite pas à leur seule intégration. Aux échelles de temps étudiées, les modèles couplés ont tendance à dériver au fur et à mesure de la prévision, et doivent donc être corrigés à la fois en cours d'intégration et après.

1.2.4.1 Correction en cours d'intégration

A l'échelle saisonnière, les modèles couplés posent plusieurs problèmes au modélisateur : celui de la dérive lente due à l'inadéquation du couplage ou des paramétrisations de certaines composantes du modèle (Delecluse et al., 1998), et celui des biais et autres erreurs du modèle. Plusieurs techniques ont été proposées pour corriger cette dérive des modèles couplés. Une solution est la **correction de flux** (Sausen et al., 1988), qui consiste à introduire un terme de flux entre l'océan et l'atmosphère au cours du couplage, en plus des flux modélisés, de manière à maintenir les valeurs moyennes des flux vus par l'atmosphère ou l'océan proches des états d'équilibre de chaque composante. Un exemple simple de correction de flux est le couplage d'anomalie proposé par Kirtman et al. (1997). Cette méthode consiste à calculer des climatologies de tension de vent de surface de la composante atmosphérique et de température de surface de l'océan de la composante océanique du modèle, puis de corriger les anomalies de tension de vent et de TSO calculées en cours d'intégration du modèle en fonction de l'écart des climatologies des composantes du modèle aux données de référence.

La réduction des erreurs des modèles dynamiques à l'échelle saisonnière a fait l'objet de nombreuses études au fil des ans. Le projet POTENTIALS (Project On Tendency Evaluations using New Techniques to Improve Atmospheric Long-term Simulations) visait à définir des méthodes d'évaluation et de réduction d'erreurs de tendance de modèles de circulation générale atmosphériques de niveaux de complexité différents. Ce projet a donné lieu à plusieurs études qui constituent la colonne vertébrale de la technique de « dynamique stochastique » développée pendant cette thèse. Guldberg et al. (2005) ont utilisé la technique de « **nudging** » ou relaxation de Newton (voir le paragraphe 2.3.1 du chapitre suivant) afin d'estimer les erreurs de tendance initiale du modèle atmosphérique ARPEGE pour un ensemble de variables du modèle. En utilisant la climatologie mensuelle de l'opposé de ces termes d'erreur comme correction systématique de la dynamique du modèle en cours d'intégration, les auteurs de l'étude montrent que les biais systématiques du modèle pendant la saison Décembre-Janvier-Février (DJF) sont considérablement réduits. Toutefois, malgré ces améliorations de l'état moyen, les prévisions saisonnières en utilisant cette méthode ne sont pas meilleures sur l'hémisphère Nord.

Dans le cadre de POTENTIALS, D'Andrea et Vautard (2000) ont proposé une autre approche à ce problème, reposant sur une méthode plus complexe avec un modèle atmosphérique global quasi-géostrophique à trois niveaux verticaux (Marshall et Molteni, 1993). Les erreurs systématiques du modèle sont évaluées en utilisant une méthode d'assimilation variationnelle,

consistant à estimer le terme de forçage nécessaire pour amener l'état du modèle vers l'état observé au pas de temps suivant. Les termes de correction ainsi définis, D'Andrea et Vautard (2000) ont corrigé le modèle en cours d'intégration en fonction de son état, caractérisé par le champ de vortacité potentielle global aux trois niveaux du modèle. Ils ont pour cela procédé à une classification du champ de vortacité potentielle sur la région Atlantique Nord-Europe. Ensuite, en cours d'intégration, la vortacité potentielle du modèle était forcée par la moyenne des K corrections correspondant aux K plus proches analogues du champ à l'instant t . Cette technique a permis d'améliorer la climatologie et la variabilité du modèle quasi-géostrophique. La représentation des régimes de temps sur la région Atlantique Nord-Europe a aussi été améliorée. D'Andrea et Vautard (2000) montrent que la dépendance des corrections de l'état du système est essentielle pour obtenir ces résultats.

Yang et Anderson (2000) ont utilisé une technique similaire à Guldberg et al. (2005) pour corriger un modèle d'océan couplé dans le but d'améliorer les prévisions de l'ENSO. Leurs résultats montrent qu'en corrigeant les erreurs de tendance initiale moyennes mais en laissant le modèle d'océan générer son cycle annuel, les prévisions sont améliorées. Barreiro et Chang (2004) ont employé une technique d'approximation linéaire des erreurs de tendance initiale afin de corriger les TSO de l'Atlantique tropical d'un modèle de circulation générale atmosphérique couplé à un modèle d'océan en couches, et ont ainsi amélioré les prévisions saisonnières de TSO et de précipitations.

Plus récemment, les travaux de Danforth et al. (2007) ont porté sur des corrections dépendant de l'état du modèle dans le but d'améliorer les prévisions à court terme. Les auteurs ont utilisé une analyse statistique des corrections d'erreurs de tendance et des anomalies prévues afin de calculer la correction du modèle en cours de prévision à chaque pas de temps par régression linéaire. Ils sont parvenus à améliorer les prévisions du modèle quasi-géostrophique de Marshall et Molteni (1993) et du modèle SPEEDY (Molteni, 2003) avec cette méthode.

Une étude de DelSole et al. (2008) a permis d'évaluer plusieurs méthodes de correction empirique d'un modèle couplé surfaces continentales - atmosphère à plusieurs échelles de temps. Les auteurs ont testé trois types de correction ne dépendant pas de l'état du modèle : le rappel des variables pronostiques vers une climatologie, la correction des erreurs de tendance initiales estimées par le nudging, et une correction à chaque pas de temps basée sur le biais mensuel moyen du modèle non corrigé. Des trois méthodes étudiées, celle basée sur la correction des erreurs de tendance initiales donne les meilleurs résultats, en utilisant les corrections moyennes du mois calendaire en cours de prévision. L'apport principal est la réduction des biais du modèle, qui permet de réduire l'erreur quadratique moyenne de la température et de l'humidité du sol ainsi que la température de l'atmosphère. Une correction linéaire dépendant de l'état du modèle (où les coefficients linéaires sont optimisés en point de grille) ne permet pas d'améliorer ces résultats.

Les différentes études mentionnées ici montrent que la correction empirique d'un modèle utilisé en prévision saisonnière peut permettre de réduire les biais et d'améliorer les prévisions, mais les résultats obtenus dépendent fortement du modèle étudié et de la méthode employée.

1.2.4.2 Correction et calibration *a posteriori*

Les sorties des prévisions d'un modèle numérique sont rarement exploitées telles quelles par l'utilisateur final. Une étape fondamentale est celle de la correction *a posteriori*. Cette cor-

rection nécessite d'étudier précisément les caractéristiques du modèle en prévision saisonnière sur une période historique, en calculant des rétro-prévisions saisonnières et en les comparant à des données de référence. Un ensemble de scores et de statistiques de ces rétro-prévisions permet d'évaluer le modèle et d'avoir une indication de la confiance qu'on peut accorder aux prévisions d'une part, et des corrections à apporter aux sorties du modèle d'autre part. Nous reviendrons plus longuement sur les méthodes d'évaluation des prévisions dans le chapitre suivant (partie 2.2).

La correction la plus simple, et qui sera faite systématiquement lors de l'évaluation des prévisions saisonnières étudiées dans cette thèse, est la **correction du biais moyen**. Les différents projets de recherche et l'implémentation de rétro-prévisions sur des périodes communes par plusieurs centres de recherche ont permis d'étudier et de développer des techniques de calibration beaucoup plus poussées, utilisant souvent une formulation probabiliste des prévisions (voir les travaux de Alessandri et al., 2011; Doblas-Reyes et al., 2005; Stephenson, 2008; Stephenson et al., 2005, pour une revue détaillée). Un des grands avantages de la formulation probabiliste des prévisions saisonnières est d'enrichir les prévisions avec une information sur les incertitudes auxquelles on est confronté à cette échelle.

Synthèse

Dans cette partie, nous avons décrit un historique de la prévision saisonnière à l'aide de modèles numériques de climat, en partant des travaux fondateurs de Manabe et al. (1965) et Shukla (1981) pour arriver aux derniers développements issus de projets internationaux récents comme ENSEMBLES.

Nous nous sommes également attardés sur des techniques de correction empirique des modèles numériques, comme celles étudiées dans le cadre du projet POTENTIALS, qui permettent de corriger des erreurs systématiques des modèles. Certaines de ces techniques s'avèrent assez efficaces, mais les résultats dépendent fortement du modèle étudié.

Un autre aspect fondamental de la prévision saisonnière à l'aide de modèles dynamiques est la prise en compte des incertitudes liées aux conditions initiales et aux choix de modélisation. Cet aspect, qui est au cœur de cette thèse, fait l'objet de la partie suivante.

1.3 Prise en compte des incertitudes dans les modèles dynamiques

Il existe deux sources majeures d'incertitudes des prévisions saisonnières avec des modèles dynamiques : les conditions initiales d'une part, et les choix de modélisation d'autre part. Dans cette partie nous abordons les différentes méthodes utilisées en prévision saisonnière pour prendre en compte ces incertitudes. Nous terminerons cette partie en présentant en plus grand détail l'ensemble des techniques couramment appelées « physique stochastique » dont nous nous sommes inspirés pour les développements réalisés au cours de cette thèse.

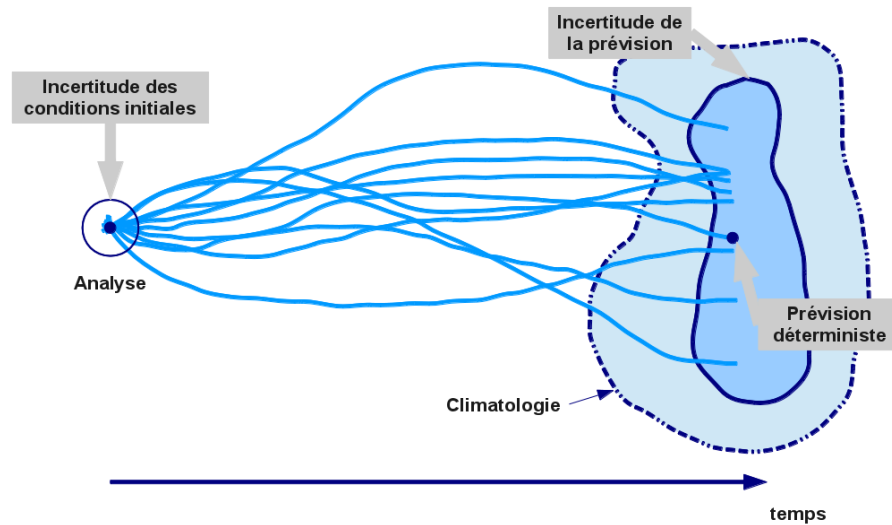


FIGURE 1.8 – Schéma représentant le principe et les avantages de la prévision d'ensemble avec les modèles de climat. (D'après Steve Easterbrook de l'Université de Toronto, « Why stochastic forecasting is better than deterministic forecasting », post de blog du 6 juillet 2010 : www.easterbrook.ca/steve/)

1.3.1 Incertitudes liées aux conditions initiales

Nous avons vu au début de ce chapitre que l'aspect chaotique de l'atmosphère rendait toute prévision déterministe de l'état de l'atmosphère impossible au-delà de quelques jours. Une petite incertitude sur les conditions initiales de l'atmosphère peut donner lieu à des prévisions assez distinctes, ce malgré des conditions aux bords identiques. Il est donc indispensable de la prendre en compte.

La formulation probabiliste des prévisions dynamiques du temps est proposée par Epstein (1969a). Il propose d'utiliser un ensemble d'équations stochastiques-dynamiques, en insérant la densité de probabilité dans les équations pronostiques du modèle. Ces travaux, suivis de ceux de Leith (1974), amènent à la **prévision d'ensemble** telle qu'elle est utilisée aujourd'hui : on implémente un nombre n de prévisions à partir de conditions initiales légèrement perturbées afin de prendre en compte l'incertitude sur celles-ci, et on forme ainsi un ensemble de n membres qui constitue la prévision. La formulation probabiliste la plus simple consiste à considérer que les membres donnent chacun une réalisation équiprobable du climat pour la saison à venir.

Le principe général de la prévision d'ensemble est présenté schématiquement dans la figure 1.8. Etant donnée une incertitude sur les conditions initiales de l'atmosphère à un temps t_0 , représentée par un cercle bleu autour des données de l'analyse utilisée, on aura une certaine incertitude de la prévision après un temps t d'intégration du modèle. Les différentes réalisations du modèle, représentées par les traits bleu clair, sont les membres de l'ensemble. Ils donnent une idée du climat après un temps t et de la densité de probabilité de chaque réalisation possible. Si on ne choisit qu'une seule réalisation comme prévision déterministe, on perd toute chance d'avoir une information sur l'incertitude de la prévision. On voit donc l'intérêt de la prévision d'ensemble pour aider à dessiner les contours des réalisations possibles sachant l'incertitude des conditions initiales.

On peut utiliser différentes méthodes d'initialisation des membres d'un ensemble. Les plus fréquemment utilisées en prévision saisonnière sont des méthodes de perturbations initiales aléatoires, indépendantes de la dynamique du modèle à l'état initial. La technique dite de **lagged-average** (Hoffman et Kalnay, 1983) consiste à partir de conditions initiales atmosphériques décalées de plusieurs heures (ou jours). Des méthodes de Monte-Carlo sont également utilisées afin de perturber les conditions initiales de manière réaliste, c'est à dire avec des structures horizontales et verticales cohérentes avec les erreurs du modèle, et des amplitudes comparables à l'incertitude estimée de l'analyse servant de condition initiale (Kalnay, 2003b).

1.3.2 Incertitudes liées aux choix de modélisation

Les différentes composantes des modèles numériques de climat sont formulées pour une certaine résolution horizontale (et verticale, le cas échéant). Les phénomènes météorologiques et physiques ayant lieu à une échelle plus fine doivent être paramétrisés. Cette résolution ainsi que la discrétisation temporelle des équations d'évolution des variables étudiées donnent lieu inévitablement à un certain nombre d'erreurs. Les incertitudes liées à la modélisation peuvent toutefois être estimées et prises en compte dans le but d'améliorer les prévisions à l'échelle saisonnière. La prévision d'ensemble à partir de conditions initiales légèrement différentes ne permet pas de quantifier ce type d'incertitude. De plus, les modèles ont tendance à se montrer trop confiants dans leur prévision, en donnant une fonction de densité de probabilité trop étroite par rapport à celle observée. Ceci n'est pas problématique en soi, à condition que le modèle ait raison dans ses prévisions. Mais à l'échelle saisonnière, c'est malheureusement loin d'être toujours le cas.

Trois types de solutions sont souvent proposées en prévision saisonnière afin de pallier à ce problème :

- la prévision multi-modèle
- les perturbations de paramètres des modèles
- les perturbations stochastiques

Nous les détaillons dans ce qui suit.

1.3.2.1 Le multi-modèle

L'approche **multi-modèle** consiste à regrouper les prévisions d'ensemble de plusieurs modèles dynamiques en un seul « super-ensemble » afin d'obtenir une estimation (assez grossière, certes) des incertitudes liées aux erreurs des modèles (Krishnamurti et al., 1999; Palmer et al., 2004). Elle a été au cœur des projets internationaux cités dans la partie précédente.

L'ensemble multi-modèle le plus simple à étudier est celui qui consiste à donner un poids équivalent à chaque membre de chaque modèle, en calculant une prévision probabiliste en supposant que chaque membre est une réalisation équiprobable, et en prenant la moyenne d'ensemble comme prévision déterministe. Malgré sa simplicité, cette approche permet en général d'améliorer les prévisions des modèles individuels. Sa force réside dans la compensation d'erreurs entre les différents modèles constituant le multi-modèle. A condition que la qualité des différents modèles inclus dans le multi-modèle soit équivalente, un modèle individuel ne sera jamais systématiquement meilleur que les autres sur toutes les régions du globe pour toutes les variables à prévoir. Ces résultats ont été montrés pour les multi-modèles des projets PROVOST (Doblas-Reyes et al., 2000) et DEMETER (Hagedorn et al., 2005). L'approche multi-modèle est

évaluée pour les précipitations africaines des rétro-prévisions du projet ENSEMBLES dans le chapitre 3. Weigel et al. (2008) ont montré que l'amélioration des prévisions probabilistes par l'approche multi-modèle résultait de l'augmentation de la dispersion d'ensemble conjointement à une réduction de l'erreur de la moyenne d'ensemble. En théorie, si les modèles avaient une dispersion suffisante et des erreurs identiques, l'approche multi-modèle ne pourrait pas battre le meilleur modèle de l'ensemble. L'approche multi-modèle est à l'heure actuelle la solution la plus pragmatique pour améliorer la dispersion d'ensemble et les performances des prévisions à l'échelle saisonnière (Weisheimer et al., 2011). Toutefois, à l'échelle mensuelle, Weisheimer et al. (2011) montrent qu'une technique de perturbations stochastiques adaptée peut obtenir de meilleurs résultats qu'un multi-modèle pour les prévisions de précipitations et de température de surface.

Forte de son succès, l'approche multi-modèle n'est plus réservée aux projets de recherche internationaux, mais est également employée en prévision saisonnière en temps réel. C'est le cas des prévisions du consortium EUROSIP (comprenant le MetOffice, le CEPMMT, Météo-France et depuis peu le NCEP), des prévisions multi-modèles américaines (le National Multi-Model Ensemble comprenant des prévisions d'ensemble de cinq centres de recherche nord-américains³) et du multi-modèle de la région Asie-Pacifique (le multi-modèle de l'APEC, (Wang et al., 2009), comprenant suivant la disponibilité des données les prévisions de jusqu'à 17 institutions différentes).

1.3.2.2 Paramètres perturbés

L'approche de **perturbation des paramètres** vise à prendre en compte les incertitudes d'un modèle liées à la formulation des paramétrisations sous-maille. L'idée de la méthode est de créer un ensemble de prévisions en modifiant les valeurs de certains paramètres empiriques. Cette méthode, proposée par Murphy et al. (2004) pour la modélisation du changement climatique, a été testée dans le modèle DePreSys du MetOffice en prévision saisonnière dans le cadre du projet ENSEMBLES, en perturbant simultanément 29 paramètres du modèle atmosphérique de manière à générer un ensemble de neuf simulations différentes. Weisheimer et al. (2011) montrent que cet ensemble donne des rétro-prévisions moins bonnes en général que les autres techniques étudiées, mais une évaluation plus fine en comparant un ensemble avec paramètres perturbés à un ensemble sans perturbation du même modèle couplé est nécessaire pour conclure quant aux mérites de cette technique. Une alternative à cette méthode est l'approche multi-physiques qui consiste à utiliser des paramétrisations différentes selon les membres de l'ensemble.

Un inconvénient majeur qu'on peut trouver à ce type de méthode est son caractère purement empirique, et la forte dépendance du modèle étudié. L'obtention d'un ensemble avec une bonne dispersion, sans perdre en qualité des prévisions, nécessite sans doute un grand nombre de tests avant de déterminer quel jeu de paramètres perturber et avec quelle amplitude. Collins et al. (2006) évoquent un grand nombre de paramètres modifiés dans le modèle du MetOffice pour obtenir une dispersion dans la réponse aux forçages anthropiques similaire à celle d'un multi-modèle. Cet effort implique un nombre d'heures de travail et de temps de calcul considérable.

3. voir le site Internet : http://www.cpc.ncep.noaa.gov/products/NMME/NMME_description.html

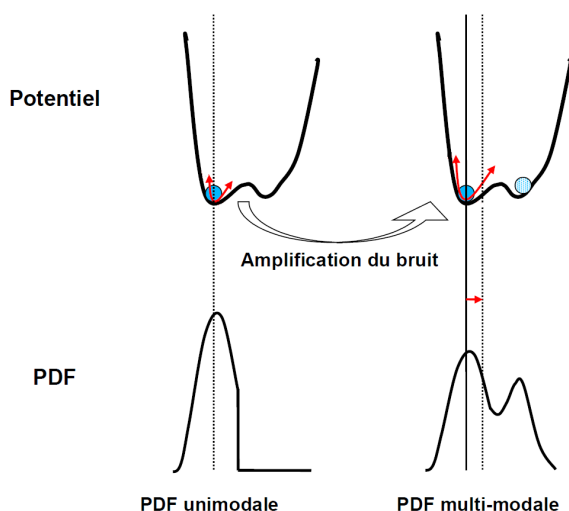


FIGURE 1.9 – Analogie du puits de potentiel expliquant l'intérêt des perturbations stochastiques dans un modèle de climat.

1.3.2.3 Physique stochastique

L'idée de **physique stochastique** est loin d'être récente : le concept est apparu au cours des années 1970. Si ce concept a fait l'objet d'études dans les décennies qui ont suivi, ce n'est que depuis une dizaine d'années qu'on voit progressivement apparaître sa mise en œuvre dans les modèles de circulation générale, en premier lieu au CEPMMT. La physique stochastique dans les modèles de climat a fait l'objet d'un « Theme Issue » de la revue *Philosophical Transactions of the Royal Society A* en 2008. Les travaux de thèse présentés dans ce mémoire reposent en partie sur cette théorie, qui fait donc l'objet d'une présentation plus détaillée dans la partie suivante.

1.3.3 Perturbations stochastiques des modèles de climat

L'idée d'introduire des termes stochastiques dans la modélisation du climat est souvent attribuée à Hasselmann (1976). Elle repose sur la séparation d'échelles de temps en deux échelles « climatiques » et « météorologiques » dans les équations différentielles utilisées dans les modèles de climat. Les échelles météorologiques représentent les échelles non résolues, qui font en général l'objet de paramétrisations caractérisant un effet moyen de ces échelles sur l'évolution des variables. Le principe introduit par K. Hasselmann consiste à prendre en compte non seulement la moyenne des effets des échelles non résolues mais également leur variance en ajoutant un terme de bruit blanc gaussien dans l'équation d'évolution. Cette méthode permet en théorie d'accroître la variabilité des modèles et de reproduire de manière plus réaliste l'influence des échelles non résolues sur l'évolution du système (Palmer, 2001).

Cette idée a été reprise sous diverses formes par des groupes de recherche en modélisation du climat mais également sous un angle plus théorique par des mathématiciens. Elle a donc donné lieu à plusieurs études d'implémentation dans des modèles de complexité différente. Le papier de Wilks (Wilks, 2008) revient sur le concept de base et sa mise en œuvre dans des modèles théoriques, notamment pour l'étude des ères glaciaires au courant des années 1980.

Une analogie souvent utilisée afin de justifier l'ajout de termes stochastiques dans les modèles est celle d'une bille dans un puits de potentiel. La figure 1.9 illustre cette analogie, décrite mathématiquement dans Palmer (2001). Le puits de potentiel en haut de la figure illustre la variabilité du climat autour de deux états, correspondant à des minima locaux de potentiel. Selon l'initialisation du modèle, la solution numérique pourrait alors converger vers un des deux états, et (dans un cas extrême) rester bloquée dans cet état pour toute la suite de l'intégration du modèle. C'est ce qui est schématisé à gauche dans la figure, où une bille se retrouve bloquée dans le premier puits de potentiel, et des petites perturbations ne lui permettent pas d'aller explorer le deuxième puits. La fonction de densité de probabilité (PDF) obtenue avec la prévision dans le cas de gauche est unimodale, et le deuxième état qui est presque aussi probable que le premier en théorie n'est pas exploré par le modèle. L'ajout de perturbations stochastiques permet d'amplifier le bruit, et donc d'imposer des perturbations plus grandes de la position de la bille ; ces perturbations permettent alors à la bille de passer d'un puits au suivant, et d'explorer les deux états. La PDF alors obtenue est multi-modale, et plus conforme à ce qui est réellement observé.

1.3.3.1 Contraintes mathématiques dans l'implémentation de termes stochastiques

L'étude des équations différentielles stochastiques est une thématique des mathématiques appliquées en plein essor. La complexité de l'approche de physique stochastique dépend du type de bruit ajouté dans les équations du modèle, en particulier du choix entre l'ajout d'un bruit additif ou multiplicatif. Dans le deuxième cas, la résolution des équations devient particulièrement délicate, et passe par des considérations mathématiques qui dépassent le cadre de cette thèse (calcul brownien, intégrales d'Itô ou de Stratanovitch, résolution de l'équation de Fokker-Planck). Les travaux de Penland et Ewald (2008) ou Majda et al. (2008) reviennent sur ces notions. L'essentiel à retenir ici est que l'utilisation de termes de bruit multiplicatif dans un modèle de climat peut avoir des conséquences désastreuses sur les sorties du modèle si les schémas numériques d'intégration des équations différentielles correspondantes ne sont pas adaptés au formalisme stochastique. On pourrait alors obtenir des résultats différents d'un modèle déterministe, mais pas pour les bonnes raisons (Penland, 2003).

La parade à ces problèmes serait d'écrire les modèles de climat entièrement sous forme stochastique. Palmer (2012) plaide en faveur du développement d'une nouvelle génération de modèles météorologiques et climatologiques formulés de manière probabiliste. Une telle solution nécessiterait un effort conjoint et considérable de nombreux chercheurs dans le domaine. En attendant, il convient d'utiliser les méthodes stochastiques dans des modèles qui n'ont pas été conçus dans ce but, mais d'interpréter avec rigueur et attention les résultats obtenus.

1.3.3.2 Les techniques de physique stochastique utilisées au CEPMMT

Les méthodes de physique stochastique mises en œuvre au CEPMMT ont fait l'objet en octobre 2009 d'un *Technical Memorandum* (Palmer et al., 2009). Les auteurs y reviennent sur les deux types de paramétrisations stochastiques testés avec le modèle de prévision d'ensemble du CEPMMT.

La première méthode de perturbations stochastiques est appelée « Stochastically Perturbed Parametrization Tendencies » ou SPPT. Cette méthode découle d'une technique développée

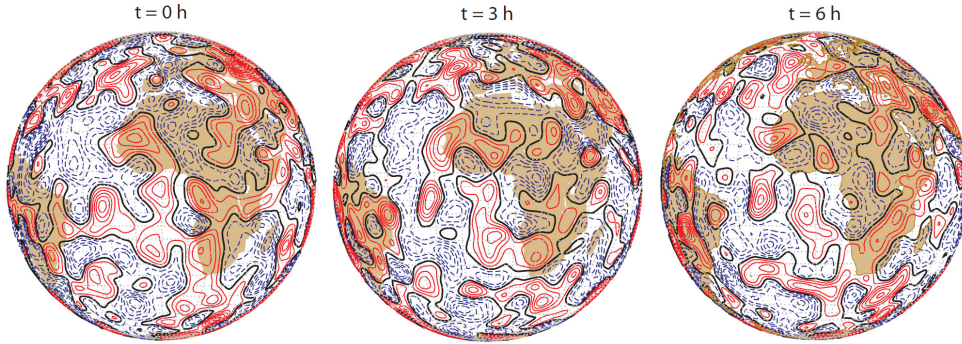


FIGURE 1.10 – Figure tirée de Palmer et al. (2009) montrant un exemple de structure spatiale et d'évolution temporelle du terme r utilisé dans le schéma SPPT.

pour le modèle de prévision à moyen terme par Buizza et al. (1999). Elle est aujourd'hui utilisée pour les prévisions opérationnelles à moyen terme ainsi que les prévisions saisonnières. Les équations du modèle peuvent s'écrire sous la forme générale suivante :

$$\frac{\partial \mathbf{X}}{\partial t} = \mathbf{A}(\mathbf{X}; t) + \mathbf{P}(\mathbf{X}; t)$$

\mathbf{A} désigne l'ensemble des processus résolus et \mathbf{P} l'ensemble des processus paramétrisés (échelles non résolues en particulier). La méthode SPPT consiste à perturber le terme de tendance \mathbf{P} pour les variables de température, d'humidité et de vitesse du vent u et v , en lui ajoutant des perturbations colinéaires aux tendances \mathbf{P} non perturbées. Ces perturbations stochastiques s'écrivent :

$$\mathbf{P}' = (1 + r\mu)\mathbf{P}$$

$\mu \in [0, 1]$ permet d'atténuer les perturbations près de la surface et dans la stratosphère (afin de supprimer certaines instabilités numériques) et r est tiré selon une loi proche d'une loi gaussienne. Les coefficients spectraux de la décomposition du terme r en harmoniques sphériques sont supposés corrélés temporellement et liés par une loi d'autorégression d'ordre 1 introduisant un bruit gaussien de moyenne nulle. Un exemple de structure spatiale du terme r et de son évolution temporelle est montré dans la figure 1.10.

En prévision saisonnière, un schéma SPPT utilisant trois échelles temporelles et spatiales pour définir les coefficients r (SPPT3) est utilisé dans le nouveau système de prévision du CEPMMT.

A partir de 2003, le CEPMMT a commencé à étudier une nouvelle approche de physique stochastique basée sur les cascades d'énergie entre les différentes échelles. Cette approche a été suggérée par les travaux de Palmer (2001). L'auteur souligne que malgré les progrès en termes de dispersion obtenus par la méthode de Buizza et al. (1999), un défaut de cette technique est son incohérence d'un point de vue énergétique (non-conservation de la quantité de mouvement ou encore de la chaleur). Partant de ce constat et d'études de pertes « numériques » d'énergie dues à l'échelle de modélisation, il propose d'utiliser un autre type de paramétrisation

stochastique qui permettrait de modéliser les incertitudes des créations et pertes d'énergie à petite échelle et leur impact sur les échelles résolues par le modèle, et d'obtenir ainsi un spectre énergétique plus réaliste.

L'algorithme CASBS (Cellular Automaton Stochastic Backscatter Scheme) décrit dans Shutts (2005) vise à corriger l'absence d'échange d'énergie cinétique entre les échelles non-résolues et les échelles résolues en forçant de manière aléatoire le champ de vorticité du modèle. Pour cela, on estime un taux de dissipation d'énergie composé de plusieurs facteurs :

- une estimation de la dissipation due à la résolution numérique
- l'effet des ondes de gravité d'origine orographique
- l'effet de la convection profonde à l'échelle sous-maille

A partir de ce taux de dissipation et d'un générateur de champs numériques aléatoires de type « Cellular Automaton » visant à reproduire la variabilité et la corrélation spatiale et temporelle de l'atmosphère à méso-échelle (Shutts, 2005), une correction du modèle est imposée en chaque point de maille par forçage de la fonction de courant. Forcer la fonction de courant plutôt que les champs de vitesse permet de maintenir la cohérence en termes de dynamique entre champs de vent et champ de pression.

L'article du « Theme Issue » de Berner et al. (2008) détaille cette méthode stochastique et ses performances en prévision saisonnière. L'utilisation de CASBS dans le modèle couplé océan-atmosphère du CEPMMT permet de réduire les erreurs systématiques du modèle pour les précipitations saisonnières sur les régions tropicales et les dérives de TSO sur le Pacifique tropical. Les scores déterministes et probabilistes calculés montrent également un gain de performances avec CASBS.

Cette méthode a également été utilisée avec le modèle du CEPMMT dans le cadre du projet ENSEMBLES. Les résultats obtenus dans le cadre du projet montrent qu'elle ne rivalise pas encore avec l'approche multi-modèle, mais confirment l'apport de CASBS aux échéances de 1 à 4 mois par rapport à l'utilisation du seul modèle du CEPMMT non perturbé (Doblas-Reyes et al., 2009).

Au lieu d'un CA pour générer des termes aléatoires, une autre forme de forçage de la fonction de courant est utilisée au CEPMMT. Le forçage de la fonction de courant est déterminé d'une manière similaire au terme r du schéma SPPT, avec une variance optimisée en fonction de la quantité d'énergie cinétique totale que l'on souhaite injecter entre deux pas de temps dans le système. Pour plus de détails on pourra se référer à Berner et al. (2009).

Les deux méthodes (SPPT3 et stochastic backscatter) sont maintenant utilisées dans le modèle du CEPMMT en prévision saisonnière. Les résultats obtenus montrent qu'elles permettent en général d'augmenter la dispersion d'ensemble et de réduire les erreurs du modèle (mais seulement pour les 3 ou 4 premiers mois de prévision). Le schéma SPPT3 a par ailleurs permis de nettement augmenter la dispersion d'ensemble des prévisions de température de surface sur la région Niño 3.4.

1.3.3.3 Autres travaux de recherche dans le domaine

Des développements similaires à ceux du CEPMMT ont vu le jour dans différents centres de recherche en météorologie et climat au cours des dix dernières années.

Au MetOffice, le système de prévision saisonnière GloSea4 (Arribas et al., 2011) comprend

deux méthodes de prise en compte des incertitudes liées à la modélisation. La première est une méthode de paramètres perturbés aléatoires décrite par Bowler et al. (2008), tandis que la deuxième est adaptée de la méthode SKEB (Shutts, 2005).

Charron et al. (2010) décrivent les méthodes de physique stochastique intégrées dans le modèle de prévision d'ensemble à moyen terme d'Environnement Canada, adaptées elles aussi des schémas employés au CEPMMT. Aux Etats-Unis, plusieurs centres de recherche intègrent des composantes stochastiques dans les paramétrisations des modèles atmosphériques de climat.

L'utilisation de techniques de perturbations stochastiques est donc devenue monnaie courante au fil des ans, et donne des résultats encourageants en termes d'augmentation de la dispersion des ensembles de prévision et de réduction des biais et erreurs quadratiques des modèles.

Synthèse

Cette partie a permis de présenter les techniques de prise en compte des incertitudes couramment utilisées pour les prévisions saisonnières. Les incertitudes des conditions initiales des simulations sont estimées en intégrant un ensemble de plusieurs prévisions partant de conditions initiales légèrement différentes. La prévision d'ensemble permet de formuler les prévisions sous forme probabiliste et de donner une information plus riche sur les états possibles du climat à l'échelle étudiée.

Afin de prendre en compte les incertitudes liées à la modélisation, trois approches sont couramment employées en prévision saisonnière : la prévision multi-modèle, la perturbation de paramètres et la physique stochastique. Dans le cadre de cette thèse, nous nous sommes focalisés sur deux de ces trois méthodes. L'approche multi-modèle a fait l'objet de plusieurs projets de recherche internationaux, et permet d'améliorer les prévisions des modèles individuels notamment grâce à la compensation de leurs erreurs systématiques (Palmer et al., 2004).

Une deuxième méthode étudiée au cours de cette thèse découle des techniques de physique stochastique. Les perturbations stochastiques des modèles de climat permettent d'accroître la variabilité des prévisions et ainsi pousser le modèle à explorer plus d'états possibles de son climat. Les techniques employées au CEPMMT présentées dans ce chapitre ont permis d'améliorer la dispersion du modèle et de diminuer les erreurs systématiques du modèle. Des travaux récents suggèrent que ces approches donnent des résultats proches de ceux d'une approche multi-modèle, voire même meilleurs à l'échelle du premier mois de prévision (Weisheimer et al., 2011).

1.4 Questions abordées dans cette thèse

Nous avons pu découvrir au fil de ce chapitre les fondements théoriques qui définissent la prévision saisonnière à l'aide de modèles de circulation générale aujourd'hui, à la fois du point de vue des sources de prévisibilité du climat, et du point de vue du développement de leur modélisation. Les récents travaux autour de la prise en compte des incertitudes liées à la modélisation nous ont incité à formuler plusieurs questions auxquelles cette thèse tente

d'apporter des éléments de réponse. Ces problématiques ont servi de jalons au cours du travail présenté ici et définissent l'organisation de la suite du manuscrit.

• *Quel est l'état de l'art actuel des modèles de prévision saisonnière ?*

Dans le cadre du projet ENSEMBLES, les rétro-prévisions sur une période de 46 ans de cinq centres de recherche ont été mises à disposition des chercheurs : l'opportunité rêvée pour évaluer l'état de l'art des modèles en prévision saisonnière. Le sujet d'étude retenu dans le chapitre 3 de cette thèse a été choisi en cohérence avec un grand effort de recherche entrepris dans le cadre du projet mondial AMMA (African Monsoon Multidisciplinary Analysis), visant à améliorer la compréhension et la prévision de la mousson africaine à toutes échelles temporelles et spatiales. Les prévisions des précipitations à l'échelle saisonnière étant un point particulièrement faible des modèles utilisés, montrer l'utilité potentielle des prévisions saisonnières du projet ENSEMBLES sur les régions de fortes précipitations en Afrique constituait un défi plus délicat à relever que la région du Pacifique tropical. Cette étude, qui a donné lieu à une publication dans la revue *Tellus A* (Batté et Déqué, 2011), permet de confirmer l'intérêt de l'approche multi-modèle sur des régions moins souvent étudiées à l'échelle saisonnière.

Cette thèse a débuté alors qu'une nouvelle version du modèle couplé servant aux projections climatiques, CNRM-CM5.1 (Voldoire et al., 2012), était développée conjointement au CNRM et au CERFACS. L'évaluation des prévisions du projet ENSEMBLES a été étendue au modèle CNRM-CM5.1, utilisé en rétro-prévision saisonnière, afin de documenter les apports et faiblesses de cette version du modèle de climat en prévision saisonnière des précipitations africaines.

• *Comment prendre en compte les incertitudes liées au modèle CNRM-CM5.1 en prévision saisonnière ?*

Dans le chapitre 4, nous présentons la méthode appelée **dynamique stochastique** développée au cours de cette thèse, qui consiste à perturber les variables pronostiques du modèle d'atmosphère ARPEGE-Climat avec des corrections aléatoires d'erreurs de tendance initiale du modèle. Cette technique vise à tirer parti des enseignements des études de correction empirique des modèles de climat (partie 1.2.4.1) et de l'ajout de perturbations stochastiques à la dynamique d'un modèle (partie 1.3.3).

• *Quelles améliorations sont théoriquement possibles en corrigeant les erreurs de la dynamique du modèle ?*

Le développement de la dynamique stochastique a donné lieu à une deuxième publication parue dans *Geophysical Research Letters* (Batté et Déqué, 2012), présentée au début du chapitre 5. Dans cette étude, une méthode de perturbations « optimales » permet de mettre en évidence les régions sur lesquelles on peut espérer améliorer les prévisions saisonnières, en adaptant la méthode de dynamique stochastique aux erreurs du modèle pour une saison donnée.

• *Existe-t-il un lien entre les erreurs du modèle et l'état dans lequel il se trouve ? Perturber le modèle en fonction de son état en cours d'intégration améliore-t-il les prévisions ?*

Nous avons décrit dans la partie 1.2.4.1 des études visant à corriger les erreurs de

tendance initiale de modèles de circulation générale en fonction de l'état du modèle (déterminé par exemple à l'aide d'analogues de circulation, voir D'Andrea et Vautard (2000)). Ces idées nous incitent à étudier l'impact d'un conditionnement des corrections d'erreurs de tendance initiale sur les performances de la méthode de dynamique stochastique définie au chapitre 4. Le chapitre 5 étudie plusieurs pistes envisagées, et détaille les résultats obtenus par les rétro-prévisions du modèle CNRM-CM5.1 employant ces méthodes.

Nous allons essayer de répondre à ces questions avec le recul que procure trois années de travail sur le sujet. Malgré tout, la plupart de ces questions sont vastes et nous avons pris le parti de les aborder sous un angle souvent restreint. Un dernier chapitre de perspectives (chapitre 6) permet de faire une synthèse générale des résultats obtenus et des pistes possibles de poursuite de ces travaux.

Avant d'aborder le vif du sujet, il est indispensable de faire un tour d'horizon des différents outils à notre disposition pour répondre à ces différentes interrogations. Le travail de thèse ayant porté sur la mise en place d'une technique de dynamique stochastique dans le modèle d'atmosphère ARPEGE-Climat, il est également essentiel de présenter ce modèle ainsi que les autres composantes du modèle couplé CNRM-CM5.1 utilisé pour les rétro-prévisions saisonnières calculées lors de ce travail. C'est l'objet du chapitre suivant.

Un point de vocabulaire... Dans ce qui suit, nous utiliserons le terme de **dynamique stochastique**, jugé moins ambigu, pour désigner la technique mise en place au cours de cette thèse, puisque celle-ci a porté sur les perturbations des variables pronostiques du modèle, à travers leur représentation spectrale, et non la partie paramétrisée représentant les processus physiques à l'échelle sous-maille.

Chapter Summary

This chapter is an introduction to seasonal predictions and an overview of the dynamical modeling techniques currently used. The first part presents the fundamentals of predictability at a seasonal time scale, and starts with a brief description of the chaos theory developed by Lorenz (1963). The weather predictability limit implies that seasonal forecasts cannot be issued in the same way as short-range weather forecasts. Seasonal forecasts are often designed to inform potential users on anomalies in surface temperature and/or precipitation means in the upcoming three-month season. Main sources of atmospheric predictability at a seasonal time scale are presented, with a focus on ENSO and Northern Hemisphere teleconnections.

The second part of the chapter is intended as an abridged history of seasonal prediction using global coupled models, from the first attempts of Manabe et al. (1965) to run a GCM over an extended period of time to the recent international research projects on seasonal prediction. A section is dedicated to model correction techniques, such as those developed and studied during the POTENTIALS project.

Then, a fundamental issue in seasonal forecasting is addressed in the third part of the chapter : taking into account initial condition and model uncertainties when designing seasonal predictions. Ensemble techniques such as multi-model forecasting and stochastic physics are described in detail. Special attention is paid to the stochastic physics techniques implemented at ECMWF.

The end of the chapter lists the main questions addressed in the rest of the thesis :

- What is the state of the art of dynamical seasonal predictions ?

This question is addressed in chapter 3, which focuses on seasonal prediction skill for precipitation forecasts over Africa for the ENSEMBLES project multi-model and the most recent version of the CNRM coupled climate model, CNRM-CM5.1. Results for the ENSEMBLES project re-forecasts were published in Tellus A (Batté and Déqué, 2011).

- How can the uncertainty related to the CNRM-CM5.1 model be taken into account at a seasonal time scale ?

Chapter 4 presents the stochastic dynamics method developed as part of this PhD thesis, which consists in perturbing the prognostic variables of the atmospheric component ARPEGE-Climat with random initial tendency error corrections.

- What are the best possible improvements one can expect using initial tendency error corrections ?

An optimal correction method was tried, drawing perturbations estimated from the tendency errors made during the actual month of the forecast in a previous run. Results are shown in a Geophysical Research Letters article (Batté and Déqué, 2012) in the beginning of chapter 5.

- Are model errors state-dependent ? Do state-dependent stochastic model perturbations improve seasonal predictions ?

Chapter 5 shows results for the implementation of several state-dependent stochastic dynamics methods in CNRM-CM5.1.

OUTILS ET MODÈLES

Ce chapitre, dont le titre aurait pu être « ingrédients et recettes » en prenant une métaphore culinaire, a pour but de présenter les différents outils dont nous disposons pour répondre aux problématiques de ce travail de thèse. Il n'a pas la prétention d'être exhaustif, mais de fournir des informations sur le modèle de climat CNRM-CM5.1 sur lequel a porté l'essentiel des travaux, ainsi que sur les différentes méthodes d'évaluation des prévisions saisonnières et les outils statistiques employés dans ce manuscrit.

2.1 Composantes du modèle couplé CNRM-CM5.1

Dans cette partie sont présentées les différentes composantes du modèle couplé du système Terre CNRM-CM5.1, utilisé en prévision saisonnière au cours de cette thèse. Cette version du modèle a été développée conjointement par le CNRM-GAME et le CERFACS en vue de l'exercice d'intercomparaison des modèles couplés CMIP5 (Coupled Model Intercomparison Project). Elle a fait l'objet d'une description plus approfondie et d'une présentation de ses principales caractéristiques dans un article de Voldoire et al. (2012). Le travail de cette thèse s'est focalisé sur la composante atmosphérique de CNRM-CM5.1, le modèle ARPEGE-Climat. Cette composante est donc présentée plus en détail dans ce qui suit, avant d'évoquer les caractéristiques des autres composantes du modèle.

2.1.1 Le modèle d'atmosphère ARPEGE-Climat

L'atmosphère est modélisée dans CNRM-CM5.1 par la version 5.2 du modèle ARPEGE-Climat, dérivé du modèle de prévision numérique ARPEGE/IFS¹ développé conjointement à Météo-France et au CEPMMT (Déqué et al., 1994).

1. pour Action de Recherche Petite Echelle Grande Echelle / Integrated Forecast System

2.1.1.1 Principes de base

Les variables de l'atmosphère qui suffisent à définir ensemble son état et son évolution, sont la température T , la vitesse horizontale du vent V_h , l'humidité q , la pression p et la densité de l'air ρ . La dynamique de l'atmosphère est régie par les équations primitives de conservation reliant ces variables :

- L'équation du mouvement horizontal (conservation du moment angulaire) :

$$\frac{d\mathbf{V}_h}{dt} = -\frac{1}{\rho}\nabla_h p - 2\boldsymbol{\Omega}_z \times \mathbf{V}_h + \mathbf{F}_h$$

- L'équation du mouvement vertical, simplifiée par l'approximation hydrostatique :

$$0 = -\frac{1}{\rho}\partial_z p - g$$

- L'équation de continuité (conservation de la masse) :

$$\frac{d\rho}{dt} = -\rho\nabla \cdot \mathbf{V}$$

- L'équation de la thermodynamique (conservation de l'énergie totale) :

$$\frac{d(c_p T)}{dt} = \frac{RT}{p} \frac{dp}{dt} + Q$$

- L'équation de conservation de la vapeur d'eau :

$$\frac{dq}{dt} = Q'$$

Q est le taux d'apport de chaleur par rayonnement, par condensation de vapeur d'eau et par convection, tandis que Q' est le taux d'apport de vapeur d'eau par condensation et évaporation.

Dans le modèle ARPEGE-Climat s'ajoute une équation de conservation de l'ozone, qui s'écrit avec un terme source S et un terme puits P :

$$\frac{dr_{O_3}}{dt} = S - P$$

L'évolution du rapport de mélange de l'ozone est évaluée par le modèle zonal de chimie MOBIDIC (Cariolle et Teyssède, 2007).

Chaque équation du modèle peut s'écrire sous une forme générique d'équation d'évolution d'une variable pronostique X :

$$\frac{dX}{dt} = \mathcal{A}(X) + \mathcal{F}(X)$$

$\mathcal{A}(X)$ englobe les effets pris en compte de manière explicite dans le modèle au vu de sa résolution, ou effets adiabatiques. $\mathcal{F}(X)$ représente l'ensemble des effets régissant l'évolution de X qui ne sont pas explicitement représentés dans le modèle, et sont donc introduits sous la forme de paramétrisations physiques intervenant à chaque pas de temps. Les variables

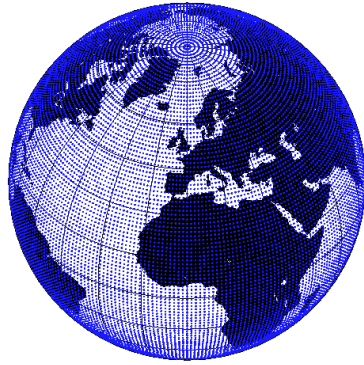


FIGURE 2.1 – Répartition sur le globe terrestre des points de la grille réduite TL127 utilisée dans le modèle ARPEGE-Climat v5.2.

pronostiques du modèle ARPEGE-Climat sont au nombre de six : la température, l’humidité spécifique, la concentration d’ozone, le logarithme de la pression de surface, la vorticité et la divergence.

Dans le modèle ARPEGE-Climat, les effets paramétrisés sont les effets radiatifs, les nuages et la turbulence, les précipitations de grande échelle, la diffusion verticale, la convection et les ondes de gravité orographiques.²

2.1.1.2 Choix de modélisation

Résolution horizontale et verticale Les coordonnées horizontales du modèle sont la longitude λ variant de 0 à 2π et le sinus de la latitude μ variant de -1 au pôle Sud à 1 au pôle Nord. Pour des raisons que nous évoquons plus loin, le modèle ARPEGE-Climat est un modèle spectral. La série des coefficients spectraux est tronquée au nombre d’onde 127. La grille linéaire réduite (Hortal et Simmons, 1991) équivalente est représentée sur la figure 2.1, et correspond à une résolution spatiale d’environ $1,4^\circ$ en latitude et longitude. Cette grille sert aux calculs des paramétrisations physiques et des termes non-linéaires du modèle, on utilise donc des transformations spectrales directes et inverses pour convertir les coordonnées des variables pronostiques entre l’espace spectral et la grille de calcul. La discrétisation en latitude permet de faire une quadrature de Gauss.

La dimension verticale est représentée dans le modèle par des niveaux hybrides entre niveaux σ et niveaux de pression, afin de prendre en compte la topographie dans les niveaux les plus bas de l’atmosphère.

Le tableau 2.1 donne les pressions et altitudes standard des 31 niveaux du modèle ARPEGE-Climat dans la configuration utilisée pour CNRM-CM5.1. La résolution du modèle est « low-top », c’est à dire que peu de niveaux sont situés dans la stratosphère. Neuf niveaux sont au-dessus de 200 hPa, tandis que six niveaux sont au-dessous de 850 hPa en pression standard.

2. Pour plus de détails, on peut se référer à la documentation en ligne du modèle ARPEGE-Climat disponible sur <http://www.cnrm.meteo.fr/gmgec/arpege-climat/ARPLI-V5.2/index.html>

Niveau	1	2	3	4	5	6	7	8	9	10	11
Pression (hPa)	7	29	50	70	90	110	132	156	181	208	238
Altitude (km)	33	24	20.5	18.5	17	15.5	14.5	13.5	12.5	11.5	10.5
Niveau	12	13	14	15	16	17	18	19	20	21	22
Pression (hPa)	270	304	340	377	417	458	500	543	588	633	679
Altitude (km)	9.86	9.08	8.32	7.60	6.90	6.22	5.58	4.96	4.36	3.79	3.25
Niveau	23	24	25	26	27	28	29	30	31		
Pression (hPa)	726	772	817	861	903	940	972	996	1009		
Altitude (m)	2727	2235	1775	1350	965	629	352	148	33		

TABLEAU 2.1 – Pression et altitude dans l’atmosphère standard des différents niveaux du modèle ARPEGE-Climat à 31 niveaux verticaux.

Modélisation spectrale La discrétisation des équations de conservation régissant l’atmosphère sur une sphère est facilitée par l’utilisation d’harmoniques sphériques et de transformées spectrales. En effet, les dérivées méridiennes et zonales sont facilitées, et peuvent être formulées de manière exacte.

La décomposition en harmoniques sphériques d’un champ A défini par ses coordonnées λ et μ s’écrit selon l’équation suivante :

$$A(\lambda, \mu) = \sum_{m=0}^N \sum_{n=m}^N A_n^m P_n^m(\mu) \cdot e^{im\lambda}$$

où $P_n^m(\mu)$ est la fonction associée de Legendre pour le nombre d’onde zonal m et le nombre d’onde global n .

$Y_n^m(\lambda, \mu) = P_n^m(\mu) \cdot e^{im\lambda}$ est solution de l’équation de Laplace :

$$\Delta Y_n^m = -\frac{n(n+1)}{a^2} Y_n^m$$

N est la troncature du modèle, sachant que la première équation n’est exacte que si N est infini. En pratique, ici, $N = 127$. Un champ spectral est donc défini à la troncature T127 par $(N+1)^2 = 16384$ coordonnées.

2.1.1.3 Discrétisation temporelle

ARPEGE-Climat utilise un schéma temporel semi-lagrangien à deux instants. On modélise la dérivée lagrangienne d’un champ X , en suivant une particule dans son mouvement, plutôt que la tendance de ce champ en un point donné (approche eulérienne).

$$\frac{DX}{Dt} = S(X)$$

Le schéma est à deux instants : $X(t + \Delta t)$ est déterminé au point de grille F à partir de $X(t)$ à un point O , représentant l’origine de la particule, et de la dérivée lagrangienne de X à un point M , représentant la position de la particule à l’instant $t + \Delta t/2$ le long de la

trajectoire. Les points O et M ne sont pas nécessairement des points de grille, et leur position doit être calculée. Le terme $S(X)$ peut être estimé par extrapolation.

$$\frac{X_F^{(t+\Delta t)} - X_O^{(t)}}{\Delta t} = S_M^{(t+\Delta t/2)}(X)$$

La méthode semi-lagrangienne nécessite plus de temps de calcul par pas de temps que la méthode eulérienne. Toutefois, le critère de stabilité est le critère de Lipschitz (les trajectoires de deux particules ne doivent pas se croiser), qui permet donc d'utiliser dans le cas de la modélisation de l'atmosphère un pas de temps plus grand que celui imposé par les conditions de Courant-Friedrichs-Lewy (CFL). Le pas de temps d'ARPEGE-Climat v5.2 à la résolution T127 est de 30 minutes.

2.1.1.4 Paramétrisations

Du fait de la résolution horizontale et verticale du modèle, certains processus physiques ne sont pas résolus à l'échelle du modèle et doivent donc être paramétrisés.

Les processus radiatifs sont paramétrisés par le schéma Rapid Radiation Transfer Model (RRTM, Mlawer et al., 1997) pour le domaine thermique et un schéma développé initialement par Fouquart et Bonnel (1980) pour le domaine solaire. RRTM calcule les transferts radiatifs sur 16 bandes spectrales, et inclut l'absorption moléculaire par les espèces majoritaires de l'atmosphère. Le schéma du domaine solaire dans ARPEGE-Climat comporte trois bandes spectrales dans le visible et l'ultraviolet, et trois bandes dans le proche infrarouge. Il inclut la diffusion de Rayleigh, l'absorption moléculaire, et les effets des aérosols troposphériques et stratosphériques (Voldoire et al., 2012).

La convection profonde est définie par le schéma de Bougeault (1985), et se déclenche avec la convergence d'humidité aux basses couches et en présence d'un profil vertical de température instable. Le schéma de Ricard et Royer (1993) est utilisé pour calculer la fraction nuageuse stratiforme, les précipitations de grande échelle et le mélange vertical turbulent.

2.1.2 Les composantes océan, glace de mer et surfaces continentales

Si le travail de cette thèse a porté sur la composante d'atmosphère du modèle couplé, nous avons vu dans le chapitre précédent l'importance des autres composantes du système terre à l'échelle saisonnière. Les différentes composantes de CNRM-CM5.1 modélisant l'océan, la glace de mer et les surfaces continentales sont présentées succinctement dans ce qui suit. Pour plus de détails, on peut se référer à Voldoire et al. (2012) et aux multiples références citées.

2.1.2.1 Le modèle d'océan NEMO

Comme l'atmosphère, l'océan vérifie des lois de conservation qui définissent son évolution à l'échelle climatique. La dynamique de l'océan est régie par trois sortes de forçages : les flux de chaleur à l'interface océan-atmosphère, qui sont fortement contrastés suivant les latitudes et impactent la couche de surface de l'océan, le forçage de la circulation de surface de l'océan par le vent (pompage d'Ekman et courants de bord Ouest), et la circulation thermohaline

induite par les variations de densité de l'eau (liées à la température et la salinité). Le rôle de la formation de la glace de mer dans les régions polaires est capital : la glace de mer étant moins salée que l'eau de mer, l'eau de surface froide et salée peut alors devenir plus dense que les eaux plus profondes de l'océan et plonger. Ces formations d'eau profonde sont équilibrées par des eaux superficielles plus chaudes qui remontent des tropiques vers les pôles. Au-delà de ces processus, la modélisation des océans est un défi en soi, du fait du tracé irrégulier des lignes de côtes. Le domaine à modéliser n'est plus une sphère régulière (comme pour l'atmosphère) mais doit être discrétisé par exemple à l'aide d'éléments finis.

Le modèle d'océan utilisé dans CNRM-CM5.1 est la composante océanique de Nucleus for European Modelling of the Ocean, ou NEMO, dans sa version 3.2 (Madec, 2008). La configuration utilisée dans CNRM-CM5.1 correspond à une grille tripolaire appelée ORCA-1°, de résolution nominale de 1° en latitude et longitude, avec une résolution accrue à 0.3° en latitude dans les tropiques. La résolution verticale est de 42 niveaux, dont 10 sur les 100 premiers mètres de profondeur. La topographie du fond de l'océan est représentée par une formulation en « partial steps ». Les équations de Navier-Stokes sont modélisées en laissant une surface libre à la surface. Le modèle comprend également les équations d'état de la température et de la salinité, pour lesquelles la diffusion turbulente et la convection sont paramétrisées.

2.1.2.2 Le modèle de glace de mer GELATO

Aux hautes latitudes, lorsque les températures de surface de l'océan deviennent inférieures au point de congélation de l'eau de mer, se forme une couche de glace. Cette couche joue un rôle sur les échanges de flux thermiques entre océan et atmosphère (en fonction de son épaisseur), sur le rayonnement solaire réfléchi par la modification de l'albédo de surface, mais aussi sur les taux de salinité de l'océan (la glace de mer étant moins salée).

Afin de modéliser ces processus, mais également la dynamique et l'évolution de la glace de mer, un modèle appelé GELATO a été développé au CNRM-GAME (Salas y Melia, 2002). La version 5 de ce modèle est utilisée dans CNRM-CM5.1.

Le modèle GELATO est un modèle multi-catégories prenant en compte l'évolution de plusieurs gammes d'épaisseur de glace par point de grille. L'advection et le transport de glace ainsi que la formation de crêtes et de feuillements sont modélisés. La grille horizontale du modèle est la même que celle du modèle d'océan de NEMO, avec une résolution d'environ 100 km. GELATO modélise également l'évolution de la couche de neige sur la glace de mer avec le schéma de Douville et al. (1995).

2.1.2.3 Modélisation de la surface et de l'hydrologie

Dans CNRM-CM5.1, c'est l'interface SURFEX (SURFace EXternalisée) qui fait le lien entre la composante ARPEGE-Climat et le schéma de surface continentale, et qui calcule les flux au-dessus de l'océan.

Le schéma de surface continentale est le modèle « Interaction between Soil Biosphere and Atmosphere » (ISBA, décrit par Noilhan et Mahfouf, 1996; Noilhan et Planton, 1989). Le modèle ISBA permet de calculer les transferts d'eau (ruissellement, drainage) et de chaleur dans le sol.

Les flux de surface sont calculés différemment selon le type de surface (végétation, lacs ou océan). Sur l'océan libre, le modèle SURFEX utilise le schéma ECUME (Exchange Coefficients from Unified Multi-campaigns Estimates, Belamari (2005)), tandis que sur la glace de mer et les lacs, les flux sont basés sur le schéma de Louis (1979). SURFEX utilise la même grille horizontale que le modèle ARPEGE-Climat, et les flux par point de grille sont estimés par pondération en fonction de la surface occupée par chaque type de surface, en se reposant sur une base de données de haute résolution appelée ECOCLIMAP.

Le ruissellement de surface et le drainage modélisés par SURFEX permettent d'alimenter le modèle de routage des fleuves « Total Runoff Integrating Pathways » (TRIP). Le modèle TRIP, développé à l'université de Tokyo par Oki et Sud (1998), assure la fermeture du cycle hydrologique à l'échelle du globe en calculant le débit d'eau douce des fleuves vers les océans. La version utilisée dans CNRM-CM5.1 a une résolution de 1° par 1° et suppose un courant uniforme et constant des fleuves à 0.5 ms^{-1} .

2.1.3 Le couplage des différentes composantes : assemblage du modèle de climat

Les différentes composantes d'un modèle de climat ne peuvent coexister sans interagir. En effet, l'évolution de chaque composante dépend directement de sorties d'une ou plusieurs autres composantes du modèle couplé. Le choix de la méthode de couplage du modèle est donc primordial.

Les composantes ARPEGE-Climat et SURFEX, les modèles NEMO et GELATO, et la composante de routage des fleuves TRIP forment trois briques du modèle. Ces différentes briques de CNRM-CM5.1 sont couplées entre elles à l'aide du coupleur OASIS version 3 développé au CERFACS (Valcke, 2006). OASIS permet d'interpoler et d'échanger les différentes données nécessaires aux briques du modèle. L'ensemble de ces données est listé en annexe de Voldoire et al. (2012). On peut citer quelques champs interpolés à titre d'exemple : la brique NEMO/GELATO envoie à ARPEGE-Climat des données relatives à l'albédo de la surface de la glace, l'étendue de la glace et les températures de surface de l'océan, tandis que le modèle ARPEGE-Climat donne à NEMO/GELATO les informations relatives aux précipitations, aux flux solaires et thermiques, ainsi qu'à la dynamique de l'atmosphère à la surface (vents zonaux et méridionaux), entre autres. Le code OASIS se charge de l'interpolation de la grille atmosphérique à la grille océanique, et inversement, en tenant compte du masque terre-mer et des différents types de surface. De même, OASIS permet l'échange de données de la brique ARPEGE-Climat/SURFEX vers TRIP et de TRIP vers NEMO/GELATO après calcul de l'écoulement des fleuves.

La figure 2.2 schématise ces différentes interactions entre les composantes du modèle couplé. La fréquence de couplage entre les différentes briques du modèle est de un jour, tandis que les modèles faisant partie d'une même brique interagissent à plus haute fréquence (chaque pas de temps généralement).

2.1.4 Implémentation d'une prévision saisonnière

Maintenant que le modèle couplé de climat CNRM-CM5.1 a été présenté, nous pouvons aborder l'implémentation d'une rétro-prévision saisonnière avec ce modèle sur une période d'étude donnée.

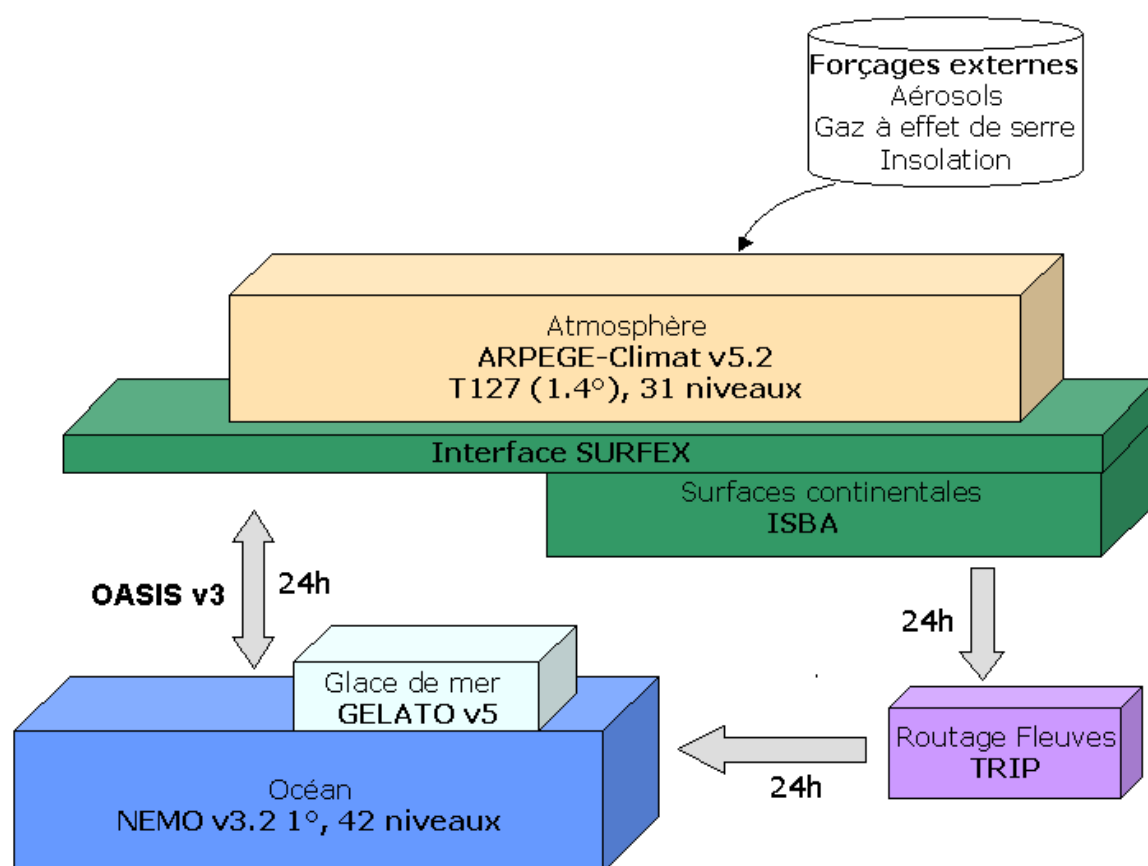


FIGURE 2.2 – Schéma des différentes composantes de CNRM-CM5.1 et des fréquences de couplage. (Source : GMGEC/ASTER, site Internet du CNRM-GAME <http://www.cnrm-game.fr/>)

2.1.4.1 Conditions initiales

Les conditions initiales du modèle doivent être précisées pour chaque composante. A cet effet, on dispose de données de réanalyses atmosphériques et océaniques (voir le paragraphe 2.2.1.2) sur une période passée. Pour la glace de mer, les données d'observation sont rares ; de même pour certains paramètres des surfaces continentales comme la couverture de neige.

Dans le cadre de cette thèse, le choix a été fait de créer des conditions initiales atmosphériques, océaniques, de surface et de glace de mer à partir d'une implémentation continue sur une période historique du modèle couplé dont l'atmosphère et l'océan seraient guidés vers des données issues de réanalyses.³ Nous reviendrons sur ce point en temps utile. Les fichiers de conditions initiales ont été enregistrés chaque mois de cette période pour les données SURFEX, GELATO, ARPEGE-Climat et NEMO, ainsi que les fichiers de conditions initiales de couplage pour OASIS.

Les prévisions saisonnières avec le modèle CNRM-CM5.1 étudiées dans le cadre de cette thèse sont initialisées à partir des mois de mai et novembre, et plus rarement de février et d'août. L'implémentation guidée du modèle sur la période historique démarre au mois de janvier de la première année, soit généralement au moins 5 mois en amont de l'initialisation de la première prévision saisonnière sur la période. Nous ferons l'hypothèse que ce délai suffit au modèle pour générer un équilibre, et donner des fichiers de conditions initiales cohérents pour chacune de ses composantes.

2.1.4.2 Choix d'implémentation

Contrairement aux simulations climatiques longues, les forçages en gaz à effet de serre, ozone et aérosols ne varient pas au cours d'une prévision saisonnière. En mode opérationnel, l'évolution des quantités de gaz à effet de serre et d'aérosols n'est pas prise en compte en temps réel dans le modèle, aussi avons nous choisi ici d'utiliser les mêmes conditions aux bords pour chaque année de la période de rétro-prévision.

Pour les aérosols, les concentrations de cinq types d'aérosols pour le mois en cours de prévision sont établies en prenant le mois correspondant et l'année 1990 dans la climatologie développée à l'IPSL (Szopa et al., 2012). Les autres forçages sont définis à partir des données du 1^{er} janvier 1990 de la simulation historique de CMIP5.

En cours de prévision, les sorties suivantes du modèle sont sauvegardées :

- les données quotidiennes :
 - de température de surface minimale
 - de température de surface maximale
 - de géopotentiel à 500 hPa
 - de pression au niveau de la mer
 - des précipitations totales
- les moyennes mensuelles :
 - de température à la surface et à 850 hPa
 - de géopotentiel à 500 hPa
 - des précipitations

3. La période couvrait 1989 à 2010, puis 1979 à 2010 dans un deuxième temps.

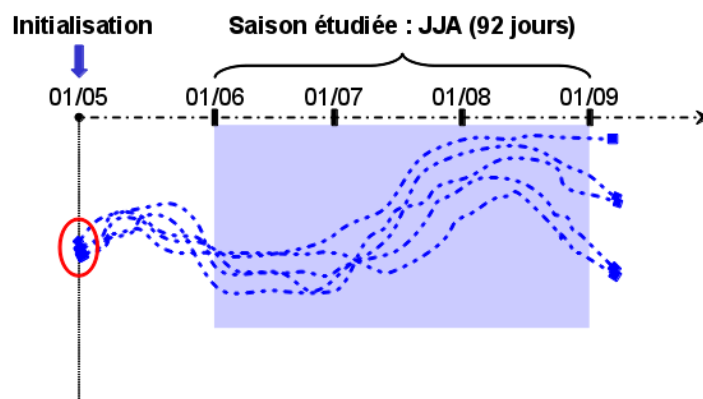


FIGURE 2.3 – Principe de la prévision saisonnière d'ensemble entre 2 et 4 mois d'échéance (cas de la prévision de la saison JJA, pour un ensemble de 5 membres).

La plupart des prévisions saisonnières étudiées dans cette thèse ont été implémentées à partir de conditions initiales de début mai et début novembre, de manière à prévoir le climat d'une saison de trois mois avec un mois d'échéance (mois 2 à 4 de la période de prévision), c'est à dire les saisons juin, juillet, août (JJA) et décembre, janvier, février (DJF).⁴ Un schéma figure 2.3 rappelle le principe de prévision d'ensemble pour une prévision saisonnière d'une saison JJA à partir de données initiales pour début mai légèrement perturbées. Dans cette thèse, le choix a été fait de privilégier la taille des ensembles de simulation plutôt que la longueur des échéances des prévisions.

Si les prévisions avec CNRM-CM5.1 permettent d'inclure les différentes composantes dans le calcul de l'évolution du climat à l'échelle saisonnière, elles ne permettent pas d'avoir recours à la méthode de lagged-average pour initialiser les prévisions d'ensemble. En effet, il faudrait pour cela enregistrer à l'échelle quotidienne les fichiers de conditions initiales de chaque composante sur la période de réanalyse, ce qui représente un coût de stockage des données considérable. La méthode retenue ici est de perturber les champs atmosphériques des conditions initiales si nécessaire.

Les prévisions saisonnières implémentées de manière opérationnelle ont beaucoup de points communs avec celles décrites ici, mais ne comportent pas à Météo-France de composantes de glace de mer et de surfaces continentales modélisées. On utilise plutôt une climatologie sur une période récente pour ces composantes. L'initialisation des prévisions est alors faite en lagged-average afin de générer un ensemble de prévisions. Dans le cadre du consortium EUROSIP, chaque mois un ensemble de 51 membres est implémenté à Météo-France. Actuellement, la version du modèle utilisée est un modèle couplé océan-atmosphère, avec le modèle ARPEGE-Climat v5.1 à la troncature T127 avec 31 niveaux verticaux, et pour l'océan la composante OPA de NEMO v3.2 avec une résolution nominale de 1°.

4. Dans ce manuscrit, pour DJF, l'année donnée correspond à celle du mois de décembre.

2.2 Evaluation des prévisions

Une fois le modèle couplé défini, il est nécessaire d'évaluer ses prévisions sur une période historique afin de dégager ses forces et faiblesses, et ainsi permettre d'affiner l'interprétation des sorties du modèle. Dans cette partie nous dressons un panorama des différentes méthodes d'évaluation des prévisions des modèles dynamiques. Ces méthodes peuvent prendre la forme de scores quantifiant une forme de distance ou de proximité entre les sorties des modèles et les données de référence, ou encore d'outils statistiques permettant d'évaluer si les modèles reproduisent ou non les liens et caractéristiques des différentes variables physiques observées. Le point commun de ces méthodes est de reposer sur des données de comparaison qui servent de jalons pour l'évaluation des modèles. Nous commencerons donc par détailler ici quelles sont ces données et comment elles sont obtenues.

2.2.1 A quoi compare-t-on les prévisions ?

2.2.1.1 Observations in-situ et données satellitaires

L'atmosphère, l'océan et les surfaces continentales font l'objet de nombreuses attentions de la part de services de recherche et d'instituts qui développent des techniques d'observation, en particulier dans le contexte actuel d'étude de l'évolution du climat liée aux forçages anthropiques. La multiplication des observations ponctuelles (lors de campagnes de mesure) ou systématiques (avec l'installation d'appareils de mesure pérennes, le lancement de nouveaux satellites, ou la systématisation de radiosondages en un point géographique donné) permet d'améliorer la connaissance du climat et de sa variabilité. Ces observations sont également une source précieuse de données de comparaison afin d'évaluer les sorties des modèles.

2.2.1.2 Réanalyses

Il est intéressant de comparer les sorties d'un modèle de climat global à des champs provenant d'observations, mais à une échelle comparable. En effet, peut-on attendre d'un modèle qui a une échelle horizontale de 1° en latitude et longitude une reproduction fidèle de données de température relevées à un point précis du globe ? La réponse est bien évidemment non, et pas seulement pour des raisons de résolution. Le développement des modèles de circulation générale est né avant tout du désir de modéliser le climat à l'échelle du globe en intégrant l'ensemble de ces informations locales. Cette technique, appelée **réanalyse**, implique d'assimiler l'ensemble des données provenant d'observations et d'en déduire via un modèle l'état des variables météorologiques à l'échelle globale sur une période historique donnée. L'avantage de cette méthode par rapport aux analyses opérationnelles faites par les centres de prévision du temps est l'homogénéité des champs sur l'ensemble de la période étudiée, du fait de l'utilisation d'une technique d'assimilation et d'un modèle de circulation générale figés et documentés.

Plusieurs centres de recherche produisent des réanalyses atmosphériques globales, comme le CEPMMT (réanalyses ERA-40 (Uppala et al., 2005) et ERA-Interim (Dee et al., 2011)), le NCEP (Saha et al., 2010), ou encore la NASA (Rienecker et al., 2011). Dans cette thèse, nous avons utilisé la base de données de réanalyses ERA-Interim du CEPMMT à la fois comme données de référence pour guider la composante ARPEGE-Climat du modèle CNRM-CM5.1

lorsque c'était nécessaire, mais aussi pour évaluer les prévisions du modèle (sauf pour les prévisions des précipitations, voir plus loin). Dans ce qui suit nous allons donc nous focaliser sur cette réanalyse en particulier.

La réanalyse est produite à partir d'une série de cycles d'assimilation de données, utilisant un modèle atmosphérique pour l'avance temporelle. Dans le cas d'ERA-Interim, le schéma d'assimilation 4D-Var est utilisé avec un cycle de 12 heures pour les variables atmosphériques. Les champs de variables atmosphériques à un niveau donné sont enregistrés toutes les 3 heures, ainsi que les données de surface. Pour les surfaces continentales, une technique plus simple reposant sur une interpolation optimale des données de surface est utilisée. Le modèle utilisé est IFS (Integrated Forecast System), avec une résolution horizontale T255, 60 niveaux verticaux et un pas de temps de 30 minutes. Les conditions de surface imposées par l'océan proviennent des bases de données océaniques du NCEP.

Les données de la réanalyse forment un jeu de données globales tridimensionnelles cohérentes entre elles et contraintes par les données d'observation utilisées dans la phase d'assimilation. En moyenne, entre 10^6 (dans les années 1980) et 10^7 (pour la dernière décennie) données d'observation sont intégrées par jour dans le schéma d'analyse 4D-Var, dont la plupart proviennent de satellites. Plus de détails sur la provenance de ces données et leur validation peuvent être trouvés dans l'article de Dee et al. (2011).

La disponibilité de données à l'échelle globale représentant l'évolution de l'atmosphère sur une période de 30 ans pourrait nous faire oublier que ces données sont calculées par un modèle. Certes, l'évolution du modèle est fortement contrainte par des observations, mais le raccourci consistant à prendre les données des réanalyses pour la réalité est dangereux et peut mener à des erreurs d'interprétation. Toutefois, ces jeux de données sont très pratiques au sens où ils permettent d'évaluer les prévisions saisonnières à échelle globale. Le fait que les différentes réanalyses soient en général concordantes entre elles, du moins sur les paramètres évalués dans le cadre de cette thèse, rassure quant à la qualité des évaluations du modèle couplé de prévision saisonnière présentées par la suite.

2.2.1.3 Le cas des précipitations

Un cas particulier de données de réanalyses parfois problématiques est celui des données des précipitations. Dee et al. (2011) évoque ce problème en détail. Les modèles de circulation générale sont généralement peu propices à représenter de manière satisfaisante les précipitations, car cette sortie de modèle dépend de plusieurs paramétrisations physiques selon qu'on s'intéresse aux précipitations stratiformes ou convectives.

On a donc choisi dans le cadre de cette thèse d'utiliser les données issues de réanalyses spécifiquement dédiées aux précipitations comme les données GPCC du Deutscher Wetterdienst (Schneider et al., 2008) ou GPCP du WCRP (Huffman et al., 2009). La réanalyse GPCC utilise les données de près de 45000 pluviomètres pour estimer les précipitations mensuelles sur une période allant de 1901 à 2007 pour la version 4 utilisée dans le chapitre 3. Ces données ne sont disponibles que sur les points de grille au-dessus des terres, à une résolution de 1° . Les réanalyses GPCP intègrent à la fois les données pluviométriques utilisées dans la réanalyse GPCC et des jeux de données de satellites pour produire des données aux échelles mensuelles et pentadaires sur une période de 1979 à 2009, et à l'échelle quotidienne sur une période restreinte à 1997–2009. Les données mensuelles utilisées pour ce manuscrit proviennent de la

version 2.2 des réanalyses, tandis que les données quotidiennes et pentadaires proviennent de la version 2.1 antérieure.

Plusieurs études comparant les différentes réanalyses des précipitations montrent que celles-ci sont fortement corrélées, avec une légère sous-estimation des précipitations dans GPCC par rapport à GPCP sur la plupart des régions.

Vu les années de disponibilité des données GPCP, les évaluations des précipitations des rétro-prévisions calculées dans le cadre de cette thèse portent le cas échéant sur des périodes se terminant en 2009, contrairement aux autres champs pour lesquels les données ERA-Interim sont disponibles sur l'ensemble de la période de rétro-prévision.

2.2.1.4 Etalons d'évaluation

Les données d'observation peuvent servir d'étalons pour l'évaluation des prévisions saisonnières. En plus de comparer directement les sorties des modèles aux observations, on peut évaluer l'apport de l'utilisation d'un modèle par rapport à des techniques statistiques reposant uniquement sur les données d'observation d'une période passée.

Dans le cas particulier des rétro-prévisions, portant sur des saisons d'une période historique, ces étalons sont calculés en « validation croisée », c'est à dire en enlevant l'année courante de rétro-prévision des jeux de données.

Un premier étalon que nous utiliserons fréquemment est la **climatologie**, désignant ici la moyenne sur les $N - 1$ autres années de la période de rétro-prévision de la variable étudiée. Pour une saison s et une année y_i de la période de rétro-prévision, la prévision saisonnière climatologique d'une variable X s'écrit :

$$X_{\text{clim}}(s, y_i) = \frac{1}{N - 1} \sum_{y \neq y_i} X_{s,y}$$

On peut également comparer les prévisions dynamiques à un étalon appelé **persistance**, qui consiste à prévoir le maintien des anomalies observées pour une variable X le mois précédent le début de la prévision (noté $M - 1$) sur l'ensemble de la saison à venir. Cette prévision s'écrit :

$$X_{\text{pers}}(s, y_i) = X_{\text{clim}}(s, y_i) + X(M - 1, y_i) - X_{\text{clim}}(M - 1, y_i)$$

A partir des données d'une période historique, des prévisions beaucoup plus sophistiquées, reposant sur des régressions de plusieurs paramètres comme par exemple les températures de surface de l'océan sur une zone précise, peuvent être calculées. Ces prévisions statistiques répondent généralement à une question précise (quelle sera l'anomalie de précipitations sur le Bénin ?) et un modèle statistique par paramètre et région est généralement nécessaire. Aussi avons-nous décidé de ne pas comparer ce type de prévisions à celles issues d'un modèle dynamique couplé, qui a pour but d'apporter une information globale sur le climat de la saison à venir.

2.2.2 Les scores d'évaluation des prévisions déterministes

Une fois déterminées les données servant à évaluer les prévisions d'ensemble des modèles couplés, il reste à définir la forme que prend cette évaluation. Les prévisions saisonnières étudiées dans cette thèse sont des prévisions d'ensemble. Ceci rend possible deux formes d'évaluation des prévisions : sous forme déterministe, en prenant la moyenne d'ensemble pour une saison s donnée comme l'unique prévision du modèle pour cette saison, et sous forme probabiliste, en regardant l'information provenant de chacun des membres de l'ensemble. Les scores d'évaluation des prévisions déterministes, appelés par abus de langage « scores déterministes » par la suite, sont présentés ici.

Le calcul de ces scores se fait selon les cas en chaque point de grille du domaine étudié, ou pour une région ou une année bien précise. Les données des prévisions sont ré-interpolées sur la grille des données de référence. On essaie pour cela de comparer les prévisions à des données d'une résolution équivalente. C'est le cas notamment pour les données ERA-Interim qui sont disponibles sur une grille de 1.5° de résolution en latitude et longitude sur le serveur du CEPMMT.

2.2.2.1 Biais

Le **bias moyen** du modèle pour une variable X et une saison s est défini comme la moyenne sur la période de rétro-prévision de l'écart entre la moyenne d'ensemble des prévisions p et les données des réanalyses o :

$$\text{Biais} = \overline{p_i - o_i}$$

où $\overline{}$ désigne la moyenne sur les années i de la période de rétro-prévision.

On peut également calculer le biais pour une année donnée. Si les biais des modèles sont des outils diagnostiques intéressants, car ils peuvent permettre d'identifier des erreurs de modélisation, ils n'entrent que rarement en considération lorsqu'on évalue une prévision saisonnière. En effet, une prévision saisonnière est souvent définie comme un ensemble d'anomalies de température, de circulation atmosphérique ou de précipitations sur la saison à venir par rapport aux normales climatiques. Ces prévisions sont faites en retirant le biais du modèle, c'est à dire en caractérisant les anomalies du modèle par rapport à sa propre climatologie plutôt que la climatologie des données de référence. Un faible biais moyen du modèle sur la période de réanalyse n'est pas une garantie d'avoir une bonne prévision. En effet, le modèle peut faire de grandes erreurs de signe opposé selon les années, et ces erreurs se compensent lors du calcul du biais moyen. Les modèles sont souvent calibrés pour minimiser les biais au cours de leur développement, par le choix de paramètres empiriques appropriés.

2.2.2.2 Erreur quadratique moyenne (RMSE)

Notons o_i les observations, c'est à dire les données de référence pour la saison ou le mois étudié l'année i de la période de rétro-prévision, et p_i la moyenne d'ensemble des prévisions du modèle. L'**erreur quadratique moyenne, ou RMSE** (pour « root mean square error »),

s'écrit :

$$\text{RMSE} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (p_i - o_i)^2}$$

o_i et p_i peuvent désigner des valeurs en un point de grille donné ou des champs 2-D des valeurs. Dans le dernier cas si on souhaite une seule valeur pour l'ensemble du domaine on peut calculer le RMSE moyen :

$$\text{RMSE} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \langle (p_i - o_i)^2 \rangle}$$

où $\langle \cdot \rangle$ désigne la moyenne spatiale (moyenne sur les points de grille du domaine étudié). Ces quantités sont le plus souvent calculées après avoir retiré le biais moyen du modèle.

Le RMSE donne une mesure de la distance moyenne sur une période donnée entre les valeurs des prévisions et les valeurs des données de référence. Il apporte donc une information plus riche que celle du biais moyen, puisqu'il n'y a plus de possibilité de compensation entre des erreurs positives et négatives du modèle vis-à-vis des données de référence. Ce score va toutefois fortement pénaliser les modèles qui produisent des anomalies plus marquées que celles observées si ceux-ci se trompent, tandis que des modèles plus « conservateurs », qui restent toujours très proches de leur climatologie, pourraient ne pas avoir le bon signe de tendance observée, mais un écart quadratique moindre avec les observations.

En prévision saisonnière, l'erreur quadratique moyenne est souvent comparée à la **dispersion d'ensemble** des modèles. Pour une taille d'ensemble n , on définit dans cette thèse la dispersion comme la déviation standard autour de la moyenne d'ensemble \hat{p} :

$$\text{Disp} = \sqrt{\frac{1}{n-1} \sum_{m=1}^n (p_m - \hat{p})^2}$$

On donne généralement la moyenne sur la période de rétro-prévision de la dispersion calculée chaque saison. Palmer et al. (2006) montrent qu'en prenant l'hypothèse d'un modèle parfaitement calibré, tel que la fonction de densité de probabilité des prévisions d'ensemble de ce modèle et de chacun de ses membres correspondent à celle des observations, le RMSE et la dispersion d'ensemble sont égaux.

Très souvent, les modèles de prévision saisonnière sont sous-dispersifs par rapport à ce critère. Nous verrons par la suite que c'est le cas pour le modèle CNRM-CM5.1 pour plusieurs paramètres. Toutefois, tendre vers ce critère d'égalité entre RMSE et dispersion ne garantit pas d'améliorer les prévisions !

2.2.2.3 Corrélation des anomalies

Un bon complément au score de RMSE est le score de corrélation, qui mesure la cohérence du modèle avec les tendances observées, sans se soucier de l'amplitude des anomalies que prévoit le modèle par rapport à celles des données de référence.

En prévision saisonnière, un score couramment utilisé est le **coefficient de corrélation des anomalies, ou ACC**, qui s'écrit comme suit pour une année i :

$$\text{ACC}(i) = \frac{\langle (\tilde{p}_i - c)(o_i - c) \rangle}{\sqrt{\langle (\tilde{p}_i - c)^2 \rangle} \cdot \sqrt{\langle (o_i - c)^2 \rangle}}$$

c désigne une climatologie de la variable étudiée, et \tilde{p} est la prévision corrigée du biais moyen. En pratique, dans cette thèse on a choisi pour une année i la climatologie $c = \bar{o}_{j \neq i}$. Déqué (1997) montre que ce score équivaut à calculer

$$\text{ACC}(i) = \frac{\langle (p_i - \bar{p})(o_i - \bar{o}) \rangle}{\sqrt{\langle (p_i - \bar{p})^2 \rangle} \cdot \sqrt{\langle (o_i - \bar{o})^2 \rangle}}$$

Le calcul de l'ACC pour chaque année i de la prévision donne une série temporelle de valeurs permettant d'évaluer l'aptitude du modèle à prévoir les anomalies spatiales de la variable étudiée. On donne souvent une valeur moyenne de l'ACC sur la période de rétro-prévision. Une alternative à la moyenne brute des ACC spatiales est le calcul de la transformée en Z de Fisher de la série de valeurs, qu'on notera zACC dans cette thèse :

$$Z(i) = \frac{1}{2} \ln \left(\frac{1 + \text{ACC}(i)}{1 - \text{ACC}(i)} \right)$$

$$z\text{ACC} = \bar{Z}$$

Un intérêt de ce score est de permettre le calcul d'intervalles de confiance indépendants de la taille de l'échantillon, la transformée en Z suivant une loi normale.

Une autre méthode pour obtenir un score agrégé est de calculer séparément la moyenne des différents moments du score d'ACC, afin d'obtenir un score appelé « mean-ACC » (Déqué et Royer, 1992), noté mACC dans cette thèse :

$$\text{mACC} = \frac{\overline{\langle (p_i - \bar{p}) \cdot (o_i - \bar{o}) \rangle}}{\sqrt{\overline{\langle (p_i - \bar{p})^2 \rangle} \cdot \overline{\langle (o_i - \bar{o})^2 \rangle}}}$$

2.2.3 Scores probabilistes

Une autre forme d'évaluation des prévisions d'ensemble consiste à tirer l'information contenue dans chaque membre de l'ensemble par l'approche probabiliste. Ce paragraphe explique le concept et les évaluations des prévisions probabilistes abordées dans ce manuscrit.

2.2.3.1 Formulation probabiliste d'une prévision d'ensemble

Soit A un événement au sens probabiliste, qui en météorologie peut prendre une forme similaire à l'assertion « les températures de la saison JJA à venir dépassent le second tercile de la climatologie » en un point géographique donné.

Chaque membre d'une prévision d'ensemble peut alors être évalué à l'aune de cet événement A , et la probabilité de l'événement A prévue par le modèle peut alors s'écrire :

$$y(A) = \frac{n_A}{n}$$

où n_A est le nombre de membres vérifiant A et n la taille de l'ensemble.

La vérification des prévisions se fait par rapport à une observation provenant des données de référence, qui vaut $o = \mathbb{1}_A$.

2.2.3.2 Score de Brier

La distance dans l'espace des probabilités des prévisions d'un événement A sur une période de rétro-prévision s'appelle le **score de Brier** (Brier, 1950) et s'écrit

$$BS = \overline{(y_i - o_i)^2}$$

Murphy (1973) et Wilks (2006) montrent qu'on peut trier les paires de prévisions et observations correspondantes selon la valeur de la probabilité prévue parmi les $J = n + 1$ possibilités $\{y_{(j)}, j \in 1, \dots, J\}$. En notant N_j le nombre de prévisions valant $y_{(j)}$, on obtient la décomposition suivante du score de Brier :

$$BS = \frac{1}{n} \sum_{j=1}^J N_j (y_{(j)} - \overline{o_{(j)}})^2 - \frac{1}{n} \sum_{j=1}^J N_j (\overline{o_{(j)}} - \bar{o})^2 + \bar{o}(1 - \bar{o})$$

$$BS = \text{Rel} - \text{Res} + \text{Unc}$$

où $\overline{o_{(j)}}$ est la fréquence relative de l'événement A lorsque la probabilité prévue vaut $y_{(j)}$, calculée en moyennant les observations correspondantes sur la période de rétro-prévision.

Le premier terme de cette décomposition, noté **Rel** pour « reliability », que nous traduirons par **fiabilité** dans ce manuscrit, est une mesure de la distance pondérée entre les probabilités prévues et la fréquence observée de l'événement ; on cherche donc un terme le plus petit possible. Cette quantité correspond au biais conditionnel des prévisions probabilistes et peut être réduite *a posteriori*.

Le deuxième terme est le terme de **résolution**, noté **Res**. Ce terme devrait être le plus élevé possible. Il mesure la capacité du modèle à séparer ses prévisions, en donnant une information utile sur la probabilité de l'événement pour la saison à venir. Si le terme de résolution est faible, cela signifie que quelle que soit la probabilité prévue, la fréquence relative de l'événement est très proche de la fréquence climatologique de l'événement, donc le modèle n'est pas d'une grande utilité.

Enfin, le troisième terme est indépendant du modèle, et donne le score de Brier d'une prévision basée sur la fréquence d'observation de l'événement. Il est noté **Unc** pour « uncertainty », ou **incertitude** en français. Dans le cas d'une prévision basée sur la fréquence d'observation de l'événement, les deux premiers termes Rel et Res sont nuls : la fiabilité de cette prévision

est maximale, vu qu'elle colle parfaitement aux fréquences observées, tandis que sa résolution est nulle.

2.2.3.3 Le RPSS

Le score de Brier permet d'évaluer la qualité des prévisions probabilistes d'un événement binaire donné. Epstein (1969b) a introduit un score généralisant le score de Brier à un ensemble d'événements probabilistes pouvant être ordonnés. Ce score est appelé « Ranked Probability Score », ou **RPS**.

Dans le cadre de cette thèse, le RPS est défini comme la moyenne des scores de Brier pour des événements $A_k = \{X \leq x_k\}$, où les x_k désignent les K -quantiles d'une variable X . Sauf mention du contraire, le RPS est calculé pour les terciles de la climatologie (en validation croisée). Toth et al. (2003) montrent qu'on peut alors écrire :

$$\text{RPS} = \frac{1}{K-1} \sum_{k=1}^{K-1} \text{BS}_k = \frac{1}{K-1} \sum_{k=1}^{K-1} \frac{1}{n} \sum_{i=1}^n (y_{k,i} - o_{k,i})^2$$

$y_{k,i}$ et $o_{k,i}$ désignent respectivement la probabilité prévue et l'observation de l'événement A_k pour l'année i de la période de rétro-prévision.

Comme le score de Brier, le RPS varie de 0 à 1. On définit un score relatif, appelé « Ranked Probability Skill Score », ou **RPSS**, en calculant les RPS du modèle étudié RPS_{mod} et de la climatologie des autres années de la période d'étude RPS_{clim} . Le RPSS est défini de manière à être positif lorsque le modèle est meilleur que la climatologie pour prévoir les différentes catégories d'événements A_k et nul lorsque les deux méthodes sont équivalentes. Ce score varie de $-\infty$ à 1.

$$\text{RPSS} = 1 - \frac{\text{RPS}_{\text{mod}}}{\text{RPS}_{\text{clim}}}$$

Comme le RPS est une moyenne de scores de Brier, on peut également utiliser la décomposition en termes de fiabilité, résolution et incertitude présentée plus haut. Le RPS de la climatologie étant égal à l'incertitude des observations, on peut réécrire le RPSS sous la forme suivante :

$$\text{RPSS} = 1 - \frac{\text{Rel} - \text{Res} + \text{Unc}}{\text{Unc}} = \frac{\text{Res} - \text{Rel}}{\text{Unc}}$$

2.2.3.4 Calibration des prévisions probabilistes

Tout comme on corrige le biais des prévisions déterministes en se servant des autres années de la période de rétro-prévision, les sorties des différents membres peuvent être recalibrées de manière à mieux correspondre aux données de référence, avant la formulation d'une prévision probabiliste. Ceci est particulièrement important pour le calcul des prévisions probabilistes basées sur des seuils climatologiques pour définir l'occurrence ou non d'un événement.

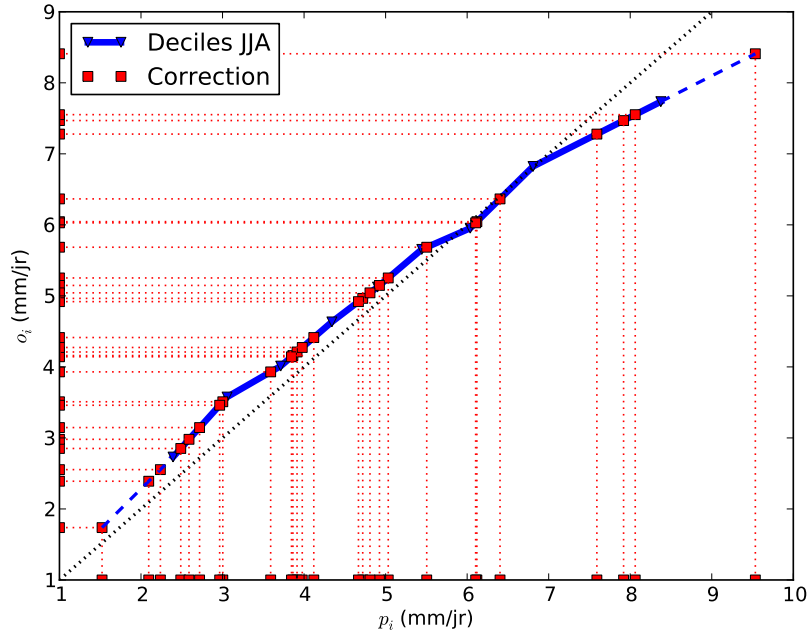


FIGURE 2.4 – Un exemple de calibration quantile-quantile des sorties des membres d’une prévision, en vue d’une prévision probabiliste (voir le texte pour plus de détails).

Dans cette thèse nous procédons lors du calcul des prévisions probabilistes à une calibration quantile-quantile des prévisions. La figure 2.4 illustre cette méthode pour un cas particulier de prévisions probabilistes des précipitations moyennes saisonnières sur un point de grille donné. Les déciles de précipitation observés sont tracés en bleu, en fonction des déciles des sorties du modèle pour les autres années de la période de rétro-prévision. Pour l’année pour laquelle on effectue la calibration quantile-quantile, les sorties des différents membres sont ensuite projetées sur cette courbe bleue (en extrapolant pour les valeurs inférieures au premier décile et supérieures au dernier décile des prévisions), afin de déterminer leur valeurs corrigées sur l’axe des ordonnées.

Dans le cas particulier de la figure 2.4, le modèle a tendance à donner des précipitations légèrement trop faibles lorsque les précipitations observées sont inférieures au cinquième décile, puis surestime les quantités de précipitations pour les saisons plus humides. Au vu de la répartition des sorties des différents membres pour la saison étudiée, la prévision probabiliste privilégie une saison assez sèche par rapport à la climatologie.

2.3 Outils statistiques

L’évaluation des prévisions ne se limite pas au calcul de scores sur des variables de sortie des modèles. On cherche également à caractériser les prévisions en calculant des indices climatiques comme ceux évoqués dans le chapitre précédent. Dans cette partie sont décrites les principales méthodes numériques et statistiques utilisées dans cette thèse, que ce soit en amont des prévisions ou en aval, pendant la phase d’évaluation des différentes versions du

modèle.

2.3.1 La technique de nudging, ou guidage des modèles

Une technique très utile pour l'étude des modèles couplés de climat est la méthode dite de **nudging**, ou relaxation de Newton, qui consiste à guider les variables pronostiques d'un modèle numérique vers des données de référence (Jeuken et al., 1996). L'équation générique du nudging s'écrit :

$$\frac{\partial X}{\partial t} = M(X) + \frac{X^{\text{ref}} - X}{\tau}$$

où $M(X)$ désigne la tendance de la variable pronostique X modélisée, X_{ref} la valeur de référence de cette variable, et τ est un temps caractéristique de rappel vers cette valeur.

La méthode de nudging a été initialement développée comme méthode empirique d'assimilation de données, de manière à inclure dans un modèle les données d'observations interpolées sur la même grille (Kalnay, 2003a). Le choix de τ est empirique et dépend de la variable pronostique nudgée. Un τ trop élevé laisse le modèle dériver loin des observations, tandis qu'un trop faible τ force le modèle à converger trop rapidement vers les données de référence, sans laisser le temps à ses composantes rapides d'ajustement de s'adapter (phase de « spin up »).

Dans cette thèse, le nudging est utilisé pour rappeler la composante ARPEGE-Climat du modèle couplé CNRM-CM5.1 vers des données issues des réanalyses, dans le but de créer des intégrations du modèle sur une période prolongée, pour disposer de conditions initiales pour les différentes composantes du modèle couplé cohérentes avec les observations historiques de cette période. Le nudging est également utilisé pour estimer les erreurs de tendance initiale du modèle, suivant une méthode proposée par Kaas et al. (1999). Nous reviendrons amplement sur ces points dans la suite du manuscrit. L'équation du nudging sur un pas de temps Δt des variables pronostiques du modèle ARPEGE-Climat s'écrit alors :

$$\mathbf{X}(t + \Delta t) = \mathbf{X}^*(t + \Delta t) + 2\Delta t \frac{\mathbf{X}^{\text{ref}}(t + \Delta t) - \mathbf{X}^*(t + \Delta t)}{\tau}$$

où $\mathbf{X}^*(t + \Delta t) = \mathbf{X}(t) + \mathbf{M}(\mathbf{X}(t), t)$

Il paraît important d'insister ici sur le fait que les prévisions saisonnières évaluées (dont on calcule des scores) dans ce manuscrit ne sont pas nudgées vers les réanalyses. La technique de nudging intervient en amont de ces prévisions.

La technique de nudging est également utilisée au cours des intégrations longues du modèle pour guider la température et la salinité de la composante NEMO vers les données de réanalyses océaniques NEMOVAR produites au CEPMMT.

2.3.2 Analyse en composantes principales et applications

Certaines caractéristiques du climat à l'échelle saisonnière ne peuvent être représentées par une simple moyenne d'un paramètre de sortie des modèles, comme la température ou les

précipitations. La technique d'analyse en composantes principales est très couramment utilisée en météorologie, sous le nom d'**analyse EOF** (pour empirical orthogonal function).

2.3.2.1 Principe

Le principe de l'analyse EOF est de simplifier un espace de grande dimension en dégagant de l'ensemble des données des directions privilégiées selon lesquelles la variance est maximale. Ces directions servent d'axes d'un espace réduit décrivant une part importante de la variabilité des données.

Afin de procéder à l'analyse EOF d'un champ géophysique sur une période donnée, on range les données dans une matrice bidimensionnelle \mathcal{M} , de manière à avoir pour chaque colonne la série temporelle des données en un point i du domaine, et pour chaque ligne le champ de valeurs au temps j . Les données sont centrées (en retirant aux champs leur moyenne temporelle sur la période d'étude).

\mathcal{M} a une dimension $n \times l$, où n est la dimension temporelle et l le nombre de points de grille. L'analyse EOF consiste à trouver les valeurs propres et vecteurs propres de la matrice de covariance de \mathcal{M} , $\mathcal{R} = {}^t\mathcal{M} \cdot \mathcal{M}$.

L'analyse EOF a donné lieu à un grand nombre d'études et de développements, et la terminologie employée varie d'une étude à l'autre. Afin de clarifier les choses, dans cette thèse on appelle EOF les vecteurs propres (de dimension l) de l'analyse en composante principale décrite plus haut, et les composantes principales (de dimension n) sont les séries temporelles des coordonnées associées à ces vecteurs. Les EOF sont orthogonales entre elles tandis que les composantes principales sont temporellement décorréliées.

En classant les valeurs propres par ordre décroissant et en retenant uniquement les N premières EOF, on peut écrire :

$$\mathcal{M}(i, j) \approx \sum_{k=1}^N \alpha_k(j) \cdot \text{EOF}_k(i)$$

où $\alpha_k(j)$ désigne la valeur à l'instant j de la $k^{\text{ième}}$ composante principale. Les valeurs propres évaluent la part de variance expliquée par le vecteur propre (EOF) correspondant.

2.3.2.2 Indices climatiques

Une application directe de l'analyse EOF est le calcul d'indices climatiques dans les données de réanalyse ou les sorties des modèles. Deux principaux indices ont été étudiés dans cette thèse.

L'indice NAO La NAO est définie dans cette thèse à l'aide de la première EOF des anomalies mensuelles de géopotential à 500 hPa des réanalyses ERA-Interim de décembre à février sur la zone Atlantique Nord-Europe. Cette zone est définie par les coordonnées de latitude comprises entre 20°N et 80°N et de longitude de 80°W à 40°E. Le champ de géopotential à 500 hPa obtenu est montré dans le chapitre précédent (figure 1.6, page 12). Un indice NAO

pour chaque saison de la période de rétro-prévision est défini en projetant les anomalies de géopotential du modèle ou des données de référence sur cette EOF.

L'indice NAM De manière similaire à l'indice NAO, l'indice du mode annulaire de l'hémisphère Nord (NAM) peut être défini à l'aide de la première EOF des moyennes zonales sur l'hémisphère Nord des anomalies mensuelles de géopotential à 500 hPa des réanalyses ERA-Interim. L'indice NAM correspond à la projection des anomalies de géopotential du modèle ou des données de référence sur cette EOF.

On notera que ces définitions ne sont pas les seules possibles pour ces indices climatiques. Nous les avons choisies pour leur simplicité de mise en œuvre, à la fois pour les prévisions d'ensemble et lors des comparaisons aux données de référence.

2.3.3 Classification automatique

La classification automatique est utilisée en climatologie afin de séparer des ensembles de champs météorologiques en groupes présentant des caractéristiques communes. Après avoir expliqué le principe de la classification automatique, nous détaillons la méthode « k-means » utilisée dans le cadre de cette thèse, et son application la plus courante qui est le calcul des régimes de temps.

2.3.3.1 Principe et mise en œuvre

Principe général L'idée de la **classification automatique** est de séparer un ensemble de points dans un espace de dimension N en k classes, ou ensembles de points possédant des caractéristiques proches. On cherche généralement à minimiser la variance intra-classe et maximiser la variance interclasses selon une norme prédéfinie.

La figure 2.5 montre de manière très schématique un exemple d'une telle classification pour un nuage bidimensionnel de points aléatoires en 4 classes différentes. La classe de points de couleur rouge peut être caractérisée par « X grand et Y petit », tandis que la classe de points de couleur cyan peut à l'inverse être définie comme la classe « X petit, Y grand ».

En météorologie ou climatologie, l'espace initial des coordonnées d'un champ géophysique est généralement de dimension très élevée. Chaque « point » dans cet espace représente le champ à un instant donné. On cherche à classer ces points de manière à regrouper les champs qui se ressemblent. La première étape est de réduire la dimension de l'espace dans lequel on travaille. Pour cela, on a souvent recours à l'analyse EOF du champ étudié. En pratique, la répartition des points dans l'espace des EOF n'est pas régulière comme celle montrée dans le cas schématique.

Une fois les métriques de distance et de séparation des classes définies, il existe plusieurs sortes de méthodes de classification. Les plus simples à mettre en œuvre sont les méthodes de classification hiérarchique et celles de centres mobiles.

Les méthodes de classification hiérarchique consistent à regrouper de manière itérative les points les plus proches entre eux. Le critère d'arrêt est basé sur le changement de la variance interclasses entre deux étapes. L'inconvénient de ces méthodes est qu'on ne connaît pas a

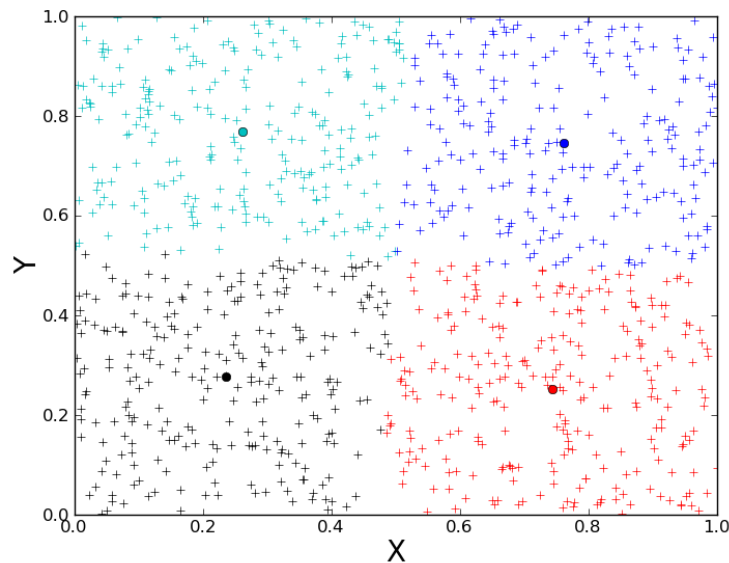


FIGURE 2.5 – Exemple de classification d'un nuage de points d'un espace $[0, 1] \times [0, 1]$ en 4 classes.

priori le nombre de classes obtenues, et que les populations des différentes classes peuvent être très variées (il est acceptable d'avoir des singletons comme classes).

Les méthodes de centres mobiles, au contraire, partent d'un nombre prédéfini de classes, et optimisent de manière itérative la répartition des points dans les différentes classes de manière à minimiser la variance intra-classe. La méthode la plus couramment utilisée est la méthode appelée « k-means » (MacQueen, 1967).

La classification « k-means » La méthode k-means est implémentée en plusieurs étapes.

- Initialisation : on tire k centres initiaux $\{C_i, i = 1..k\}$ parmi la population de points à classer
- Itération :
 1. on regroupe les points dans k classes telles que chaque classe i contienne les points plus proches de C_i que des autres centres C_j ($j \neq i$).
 2. on calcule les barycentres des classes obtenues (pas forcément des points de la population) et on remplace les centres $\{C_i\}$ par ces barycentres.
 3. on calcule la variance intra-classe totale de la classification.
- Terminaison : on répète les étapes de l'itération jusqu'à ce que l'itération supplémentaire ne réduise plus la variance intra-classe de la population (critère d'arrêt basé sur un seuil).

Plusieurs améliorations ont été suggérées pour cette méthode. Elles portent pour la plupart sur la méthode d'initialisation afin de pallier à un problème récurrent de la technique k-means : la solution obtenue peut être un minimum local de la variance intra-classe sans avoir convergé vers le minimum global de la population de points, suivant les points initiaux choisis.

Dans le chapitre 5, la méthode utilisée, appelée « k-means ++ », consiste à implémenter la méthode k-means plusieurs fois et de retenir la solution vérifiant le mieux un critère donné. Nous reviendrons sur ce point en temps voulu.

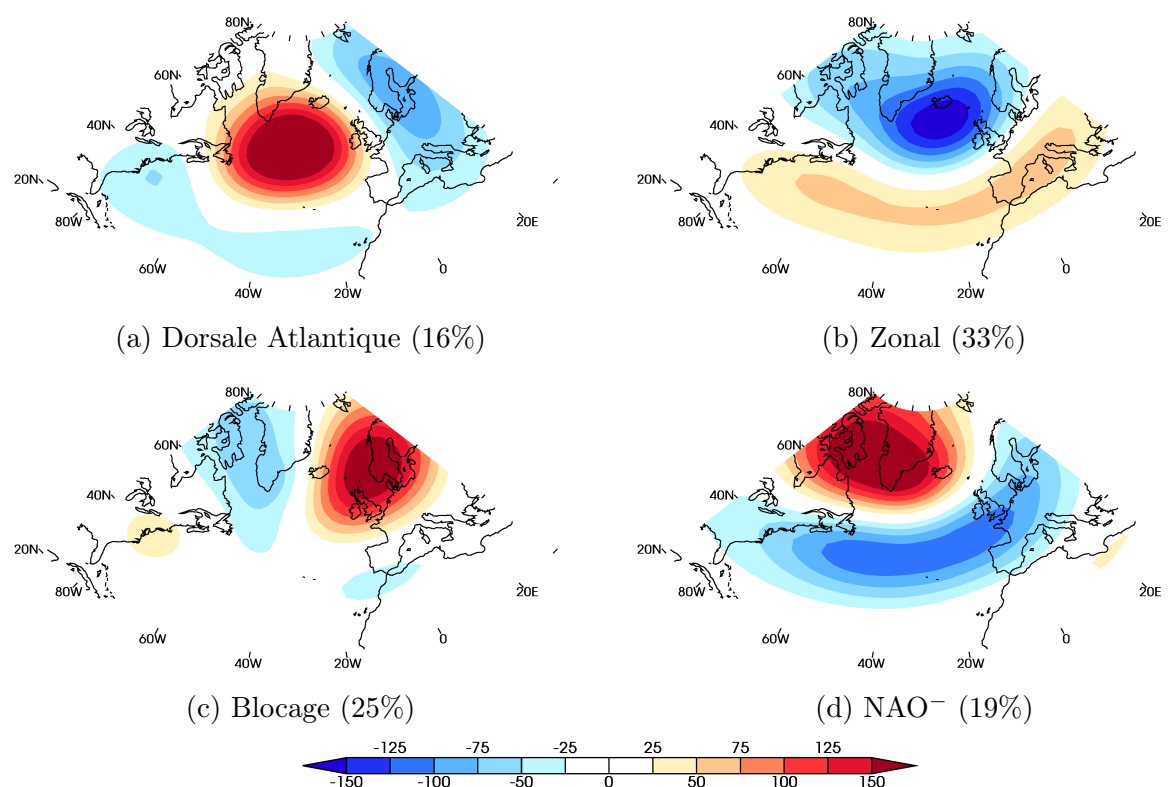


FIGURE 2.6 – Régimes de temps sur la région Atlantique Nord–Europe pour le géopotential à 500 hPa pour la saison DJF 1979–2010, calculés à partir des données quotidiennes des réanalyses ERA-Interim. Les quatre figures montrent les anomalies (en m) des centres des quatre classes obtenues par une classification k-means, leur nom couramment employé en météorologie, et leur fréquence calculée avec un seuil de définition du régime de trois jours.

2.3.3.2 Classification en régimes de temps

L'application la plus courante de la classification automatique en sciences du climat est l'étude des régimes de temps. On appelle régimes de temps des états préférentiels de la circulation atmosphérique sur une région donnée, à l'échelle intra-saisonnière. Les transitions entre différents régimes résultent de l'influence de phénomènes à l'échelle synoptique sur la circulation de grande échelle.

L'étude des régimes de temps a connu son essor au cours des années 1980-1990, avec la caractérisation de régimes de circulation dans les données des réanalyses atmosphériques (voir par exemple Cheng et Wallace, 1993; Kimoto et Ghil, 1993a,b; Michelangeli et al., 1995). Plusieurs études à l'aide de modèles quasi-géostrophiques (Marshall et Molteni, 1993) ou modèles de circulation générale (par exemple Corti et al., 2003; Hansen et Sutera, 1990) trouvent des régimes similaires à ceux des données de réanalyses.

Toutefois, l'existence de ces régimes est parfois contestée. Les travaux de Stephenson et al. (2004) et Christiansen (2007) remettent en question la significativité des régimes obtenus.

Dans cette thèse nous nous intéresserons aux régimes de temps de la région Atlantique Nord–Europe au chapitre 4, afin de caractériser la représentation des propriétés de ces régimes

(durée, fréquence) par CNRM-CM5.1 en prévision saisonnière, et l'impact de l'introduction de perturbations stochastiques dans le modèle. Il est généralement admis qu'un nombre de quatre régimes pour la saison d'hiver (DJF) donne des centres stables.

En suivant une méthode proche de celle de Michelangeli et al. (1995), on peut déterminer ces régimes pour les données d'anomalies de géopotentiel à 500 hPa des réanalyses ERA-Interim, pour les hivers de la période 1979–2010. Les données sont projetées sur la base des premières EOF de manière à expliquer plus de 80% de la variance. L'étape suivante est la classification k-means des données dans l'espace des EOF, en cherchant $k = 4$ classes. Un seuil de 3 jours est fixé comme durée minimale d'un régime. Les jours correspondant à une transition entre régimes ne sont pas attribués. Les régimes de temps ainsi obtenus et leur fréquence sur la période sont montrés dans la figure 2.6.

Nous utiliserons également la classification automatique au chapitre 5 afin d'étudier le conditionnement des perturbations du modèle ARPEGE-Climat en fonction de l'état du système.

2.4 Synthèse

Dans ce chapitre ont été présentés les outils numériques et statistiques utilisés dans le cadre de cette thèse. Les principales caractéristiques du modèle couplé CNRM-CM5.1, développées de manière plus détaillée dans l'article de Voldoire et al. (2012) et les références qui y sont citées, ont été rappelées.

Nous avons également abordé l'évaluation des prévisions des modèles de climat en cherchant à répondre à deux questions :

1. A quelles données de référence peut-on comparer les sorties d'un modèle numérique de climat ?
2. Quels sont les outils statistiques d'évaluation des modèles ?

L'accent a été mis sur les réanalyses, en particulier celle du CEPMMT, ERA-Interim, qui sert de référence pour ce travail de thèse. Nous avons détaillé les différents scores de prévisions déterministes et probabilistes qui serviront à évaluer les sorties des modèles.

Enfin, la dernière partie de ce chapitre a permis de décrire quelles étaient les méthodes statistiques utilisées dans le cadre de cette thèse, pour calculer des indices climatiques ou caractériser la variabilité à l'échelle intra-saisonnière, dans les modèles ou les données de référence.

Chapter Summary

This chapter presents the modelization and statistical tools used during this PhD thesis. Main characteristics of the CNRM-CM5.1 global coupled model, examined in more detail in Voldoire et al. (2012) and references therein, are presented in the first part of the chapter.

The evaluation methods for climate model seasonal forecasts used in this work are explained in detail in the second part of the chapter, while answering two main questions :

- 1. What reference data can outputs from a numerical climate model be compared to ?*
- 2. Which statistical tools are used to assess model performance ?*

The focus is set on reanalysis datasets, such as ERA-Interim developed at ECMWF and used as a reference in this work. Several deterministic and probabilistic skill scores used to evaluate seasonal re-forecasts are described in detail.

Last but not least, statistical methods used to compute climate indices, or characterize intra-seasonal variability in models or reference datasets, are presented in the third part of the chapter.

DEUXIÈME PARTIE

L'APPROCHE MULTI-MODÈLE

PRÉVISIONS D'ENSEMBLE DES PRÉCIPITATIONS SUR L'AFRIQUE

Sur une grande partie du continent africain, l'eau est une denrée précieuse et indispensable pour les populations locales. Plus de 40% de la population d'Afrique habite des zones arides, semi-arides ou sèches, et les projections démographiques estiment que la population du continent pourrait doubler d'ici à 2050. De nombreuses zones semi-arides sont déjà densément peuplées. Par ailleurs, les populations qui vivent de l'agriculture ont rarement les moyens de s'adapter aux aléas du climat. La plupart du temps, l'eau qu'elles utilisent pour leurs cultures est issue de précipitations intenses, mais concentrées sur une courte période de l'année. Un récent rapport du Programme des Nations Unies pour l'Environnement montre l'inégalité entre les différents pays d'Afrique face à la question de l'accès à l'eau (UNEP, 2010). Dans ce contexte, les prévisions des précipitations aux échelles intra-saisonnière et saisonnière sont d'un grand intérêt pour les décideurs locaux.

L'Afrique de l'ouest, sujette à une mousson dont les caractéristiques ont évolué au cours du siècle dernier, a fait l'objet de nombreux efforts de recherche, afin d'améliorer la compréhension et la prévision de la mousson aux échelles synoptiques à climatiques. Le projet AMMA (Analyses Multidisciplinaire de la Mousson Africaine) a pour but de coordonner des études multi-échelles du phénomène, mais également d'étudier les applications possibles d'une meilleure prévision de la mousson (gestion de l'eau, santé publique). La prévision saisonnière de la mousson africaine est également au cœur du forum PRESAO (PREvisions Saisonnières en Afrique de l'Ouest, Hamatan et al. (2004)), mis en place en 1998, qui fournit chaque année des prévisions statistiques et dynamiques des précipitations et débits de cours d'eau pour la saison de mousson.

La première partie de ce chapitre est constituée d'un article paru dans la revue *Tellus A* évaluant les prévisions saisonnières de précipitations sur l'Afrique des modèles du projet européen ENSEMBLES. L'étude est ensuite étendue dans une deuxième partie à deux régions distinctes de l'Afrique de l'ouest, le Sahel et le Golfe de Guinée. La troisième partie de ce chapitre s'attache à décrire les différences entre les prévisions de la version du modèle couplé du CNRM utilisé dans ENSEMBLES, et la nouvelle version CNRM-CM5.1. Nous regarderons aussi l'impact d'une méthode de dynamique stochastique, détaillée au chapitre 4 de ce mémoire

de thèse, sur les résultats.

3.1 Rétro-prévisions du projet ENSEMBLES des précipitations africaines

3.1.1 Résumé de l'article publié dans *Tellus*, Series A

Dans le cadre du projet européen ENSEMBLES de la Commission Européenne, plusieurs jeux de rétro-prévisions saisonnières ont été produits pour la période 1960–2005. Les prévisions de chacun des cinq centres de recherches participant à cet exercice (CEPMMT, IFM-Geomar, Météo-France, UK MetOffice et CMCC-INGV) étaient initialisées pour février, mai, août et novembre de chaque année de la période.

L'objectif de cet article, publié en mars 2011 dans la revue *Tellus A*, est d'évaluer les performances des prévisions saisonnières des cumuls de précipitations sur des régions clés de l'Afrique au cours de saisons caractéristiques. On s'est attaché à l'étude de trois régions et des quatre saisons correspondant aux échéances de deux à quatre mois après initialisation de la prévision, à savoir :

- L'Afrique de l'ouest au cours du début et du pic de la mousson, en juin-juillet-août (JJA) ;
- L'Afrique du sud pendant l'été austral (décembre-janvier-février, DJF), au cours duquel les précipitations peuvent dépasser en moyenne les 10 millimètres par jour sur Madagascar ;
- La Corne de l'Afrique pendant les deux saisons transitoires des « long rains » (mars-avril-mai, MAM) et des « short rains » (septembre-octobre-novembre, SON), vitales pour la région, comme l'ont montré les famines consécutives à la sécheresse sans précédent enregistrée en 2011.

L'étude vise à illustrer et quantifier les apports de l'approche multi-modèle dans les prévisions de précipitations sur ces régions. Afin de couvrir l'ensemble de la période des rétro-prévisions, on a comparé les prévisions aux données de pluviométrie réanalysées de GPCC. L'évaluation des prévisions sous forme déterministe et probabiliste utilise l'éventail de méthodes présenté dans le chapitre précédent. A cela s'ajoute un calcul idéalisé de valeur économique potentielle des prévisions, pour lequel on compare à l'aide d'un modèle coût-perte très schématique les économies ou dépenses réalisées en utilisant le modèle ou multi-modèle, la distribution climatologique, ou un modèle « parfait » reproduisant la probabilité observée pour un événement probabiliste donné (voir le paragraphe 3.1.2.4).

Les conclusions de l'article sont contrastées selon les régions d'étude et les méthodes d'évaluation. On montre que l'utilisation d'un multi-modèle permet d'améliorer la dispersion d'ensemble sans accroître l'erreur quadratique moyenne des prévisions des cumuls de pluie, et qu'en matière de corrélation d'anomalies, le multi-modèle permet en général d'accroître le score des modèles individuels. Dans le cas des scores probabilistes, les résultats sont plus probants sur l'Afrique de l'ouest et l'Afrique du sud. De manière générale, les travaux présentés confirment l'importance de la bonne représentation dans les modèles couplés des liens entre la TSO et les anomalies de précipitations, puisque les meilleurs résultats sont obtenus au-dessus des zones les plus influencées par ces conditions océaniques, et pour lesquelles le multi-modèle parvient à reproduire les structures de covariance d'anomalies de précipitations avec les TSO.

L'article vient compléter des études précédentes portant sur les prévisions des précipitations africaines du premier jeu de rétro-prévisions saisonnières du projet ENSEMBLES (stream 1) comme Philippon et al. (2010) et Bouali (2009). Son originalité est de présenter l'évaluation des prévisions saisonnières sur une période plus longue que la plupart des études, et en utilisant uniquement des post-traitements simples et robustes (correction des biais et calibration quantile-quantile pour les prévisions probabilistes).

3.1.2 Article : Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models : skill of the ENSEMBLES project multi-model ensemble forecasts

By Lauriane BATTÉ and Michel DÉQUÉ *Météo France, CNRM-GAME, 42 avenue G. Coriolis, FR-31057 Toulouse Cedex, France*

(Manuscript received 22 June 2010; in final form 28 October 2010)

Citation : L. Batté and M. Déqué (2011). Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models : skill of the ENSEMBLES project multimodel ensemble forecasts. *Tellus*, 63A : 283–299. doi : 10.1111/j.1600-0870.2010.00493.x.

3.1.2.1 Abstract

ENSEMBLES stream 2 single-model and multi-model ensemble seasonal precipitation forecasts are evaluated over the African continent with respect to GPCC precipitation data for the 1960–2005 time period using deterministic and probabilistic skill scores. Focus is set on three regions of Africa during main precipitation seasons : West Africa during boreal summer, southern Africa during austral summer, and the Greater Horn of Africa during the “long rains” and “short rains” transition seasons. The 45-member multi-model improves the ensemble spread-skill ratio over all regions, which translates into enhanced skill in terms of anomaly correlation and ranked probability skill scores for climatological precipitation deciles over West Africa and southern Africa. Results are contrasted depending on the region, and probabilistic formulations of the ensemble predictions after a quantile-quantile calibration give valuable information essentially over areas where deterministic skill is found. Probabilistic skill scores illustrate the range of possibilities for more user-related applications of ensemble seasonal forecasts. A simple illustration using a cost-loss model shows that model potential economic values can reach over 10% depending on the regions and occurrences studied.

3.1.2.2 Introduction

Multi-model ensemble forecasting is now a mainstream method used for medium-range to seasonal predictions as well as climate modeling. This method addresses two sources of uncertainty in dynamical models : ensemble forecasting accounts for uncertainties due to initial conditions while the use of several models with different physical formulations and parameterizations helps widening the ensemble spread and estimating single-model uncertainties (Palmer et al., 2004).

Seasonal-to-annual predictions from five state-of-the-art European global coupled climate models were run as part of the EU-funded ENSEMBLES project. Recent assessment results for

different variables (sea-surface temperature, near-surface temperature, 500 hPa geopotential height, precipitation) in Doblas-Reyes et al. (2009) show that multi-model ensemble seasonal predictions (at least up to 4 month lead) have higher skill than other methods used to reduce model uncertainties such as perturbed-parameter (Collins et al., 2006) or stochastic-physics (Berner et al., 2008) approaches. These alternative methods will therefore not be discussed in this paper.

The focus in this paper is set on seasonal precipitation forecasts over the African continent, and more specifically over sub-regions with noteworthy seasonal precipitation cycles. These precipitation forecasts are of interest in regions subject to persisting drought if used as early warning and decision-making tools in water supply management. At a seasonal time scale, the chaotic nature of atmospheric dynamics make it difficult for GCM precipitation forecasts to provide chronological information and the long-term predictability relies mostly on ocean conditions (Shukla, 1998). As such, GCM precipitation outputs are not used for prediction of monsoon onset dates without further analysis of other fields, such as circulation patterns. Quantitative precipitation seasonal forecasts show little skill when compared to persistence or climatology, and extracting useful information from such forecasts often requires elaborate post-processing methods such as statistical adaptation (Bouali et al., 2008; Garric et al., 2002). These research efforts were led as a consequence of poor seasonal precipitation forecasting skills in past ensemble forecast experiments. Another way of dealing with ensemble forecast information is to use a probabilistic formulation. This approach avoids quantitative forecast issues by yielding probabilities for different possible outcomes of a given event (Déqué and Royer, 1992). The evaluation of such forecasts is therefore more user-related, in the sense that some probabilistic scores can be highly dependent on the model actual usefulness in a given situation. Evaluation of the ENSEMBLES stream 1 (1990–2001) seasonal forecasts skill over West Africa during the monsoon season showed that direct model precipitation outputs can be improved using model-output statistics (MOS) based on simulated atmospheric dynamics (Philippon et al., 2010). Our goal here is to evaluate seasonal forecast skill for the ENSEMBLES stream 2 monthly ensemble re-forecasts over the extended 1960–2005 period using various deterministic and probabilistic scores, focusing on the precipitation outputs of the models. Additional analysis of precipitation forecast skill related to sea-surface temperatures (SSTs) is also provided.

This paper is organized as follows. The ENSEMBLES stream 2 forecasts and comparison datasets are described in the following section, along with precipitation climatologies over the sub-regions studied in this paper. Section 3 details the methodology and skill scores calculated throughout our study. Main results are presented in section 4, and section 5 summarizes general conclusions.

3.1.2.3 Data and regions of interest

ENSEMBLES project seasonal re-forecasts The ENSEMBLES seasonal re-forecasts used in this study are from the Stream 2 experiment. In this experiment five coupled global circulation ensemble models were run. Participating research centers are the European Centre for Medium-Range Weather Forecasts (ECMWF), the Leibniz Institute of Marine Sciences at Kiel University (IFM-GEOMAR), Météo France (MF), the UK Met Office (UKMO) and the Euro-Mediterranean Centre for Climate Change in Bologna (CMCC-INGV). Table 3.1 shows basic characteristics of each model. Each ensemble comprises nine runs initialized with different

TABLEAU 3.1 – Basic characteristics of the main components of the five ENSEMBLES stream 2 models used for seasonal forecasting

Centre	Atmosphere component	Ocean component
ECMWF	IFS CY31R1 T159/L62	HOPE 0.3°-1.4°/L62
IFM-Geomar	ECHAM5 T63/L31	MPI-OMI 1.5°/L40
Météo France	ARPEGE4.6 T63/L31	OPA8.2 2°/L31
UKMO	HadGEM2-A N96/L38	HadGEM2-O 0.33° – 1°
CMCC-INGV	ECHAM5 T63/L19	OPA8.2 2°/L31

sets of ocean reanalyses generated from wind stress and SST perturbations. The full multi-model ensemble (MME) is therefore made of 45 members. Further details may be found in Weisheimer et al. (2009).

Stream 2 seasonal re-forecasts were run from February, May, August and November over the 1960–2005 period. Seasonal forecasts for each month 2-4 average (March-April-May or MAM, June-July-August or JJA, September-October-November or SON and December-January-February or DJF) were used in this study. The multi-model ensemble combination method used for deterministic forecasts is a simple 45-ensemble mean, without individual model weighting.

Data were retrieved from a ECMWF server.¹ The precipitation fields are available for each model on a global $2.5^\circ \times 2.5^\circ$ grid.

Reference data : the GPCC monthly global dataset The Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis version 4 dataset for monthly mean precipitation was used as the reference dataset for climatology and verification of the ENSEMBLES stream 2 hindcasts. This dataset, described in Schneider et al. (2008), is a reanalysis of monthly rain-gauge data from up to 45000 land stations around the globe using GPCC climatology as analysis background, and covers the 1901–2007 period. It is the most appropriate GPCC product for hydrometeorological model verification. The choice of GPCC data was also based on station availability over the regions of interest and during the entire ENSEMBLES stream 2 period (1960–2005). The $1^\circ \times 1^\circ$ gridded data was used, so the ENSEMBLES predictions were regridded onto the corresponding grid for comparison purposes.

African sub-regions Sub-regions of the African continent were defined so as to focus on areas subjected to significant seasonal precipitation amounts. Figure 3.1 shows GPCC mean precipitation from 1960 to 2005 over the African continent for the four three-month seasons defined earlier. The impact of sea-surface temperature (SST) over the ENSO region and the tropical Atlantic on African rainfall is thoroughly discussed in Camberlin et al. (2001). Other teleconnections between East-African rainfall and Indian Ocean SST have also been demonstrated (Bowden and Semazzi, 2007).

1. http://www.ecmwf.int/research/EU_projects/ENSEMBLES/data/data_dissemination.html

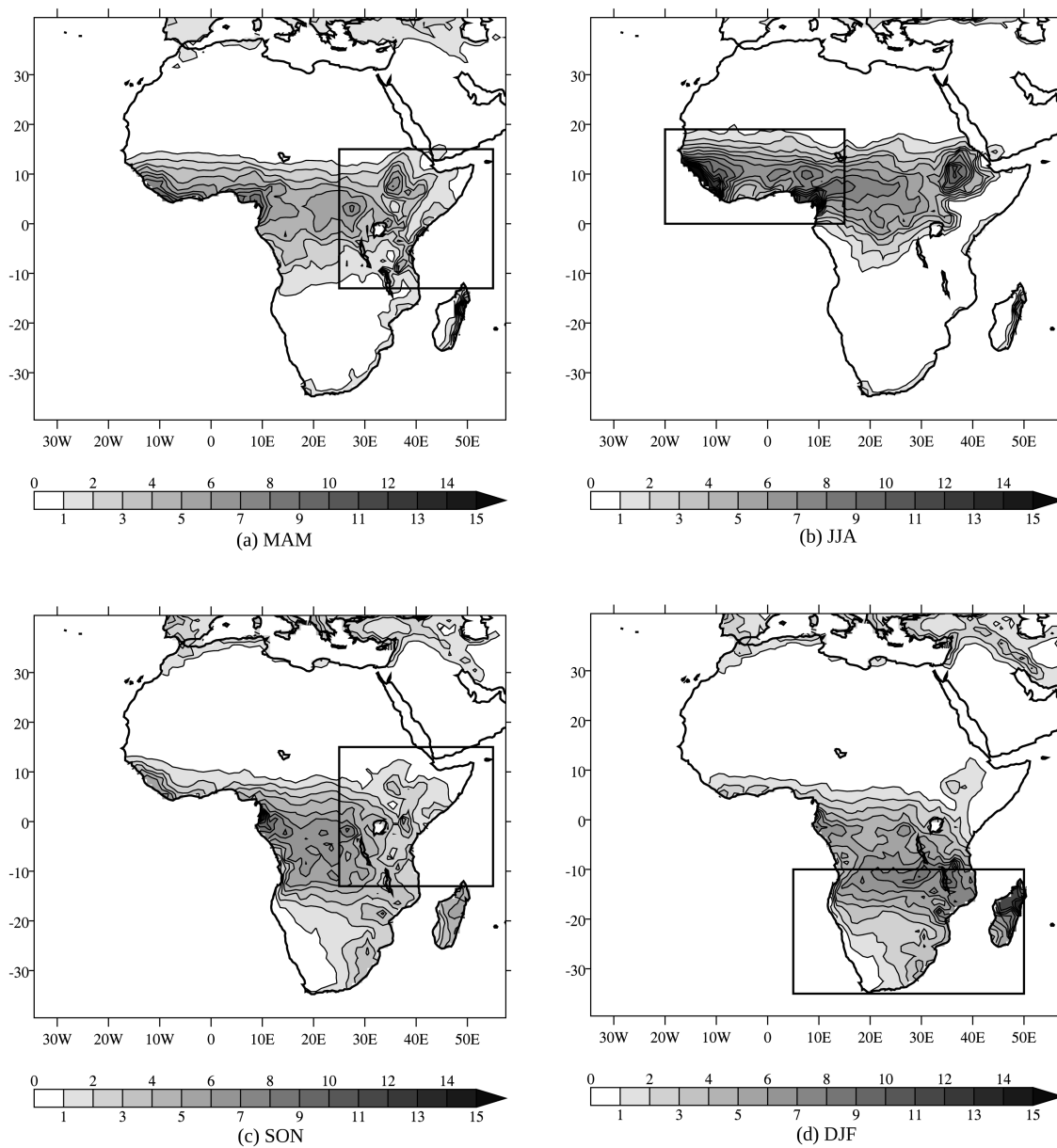


FIGURE 3.1 – GPCC seasonal mean precipitation (mm per day) over the African continent for four three-month seasons, averaged over the 1960–2005 period. Boxes show the regions of interest for the given seasons

The quantitatively most important precipitation event over the African continent is the West African Monsoon. High mean daily precipitation rates are found for GPCC data during the JJA season along latitudes ranging from approximately 2° to 15° North, with a peak locally exceeding 16 mm per day over Sierra Leone.

Other areas of Africa subjected to seasonal rainfall phenomena are the Greater Horn of Africa during the MAM and SON inter-seasons, and southern Africa during austral summer (DJF). More details on the geographical bounds of these regions in this study and the meteorological phenomena observed there are given below.

- *The West African region*

Although JJA rainfall affects most of the longitudinal extent of the African continent, the region used in this study is restricted to West Africa, ranging from the Gulf of Guinea and the Bight of Biafra to the Atlantic coast alongside Senegal. Grid points of the domain referred to hereafter as “West Africa” (WA) cover latitudes from 0° to 19° N and longitudes from 20° W to 15° E.

Seasonal means cannot illustrate the spatial shift of precipitation during the West-African Monsoon, however comparison between the MAM, JJA and SON precipitation rates over the sub-sahelian regions as well as a clear latitudinal gradient of precipitation rates are consistent with the northward shift of the ITCZ during the monsoon season documented by Sultan et al. (2003). Several studies have shown evidence of strong positive correlation between Gulf of Guinea precipitation and tropical South Atlantic SSTs as well as negative correlation between Sahelian precipitation and Niño 3 SSTs during July, August and September (Camberlin et al., 2001; Giannini et al., 2005; Vizzy and Cook, 2001).

Figure 3.2 (a) shows a preliminary evaluation of how the ENSEMBLES MME ensemble mean replicates inter-annual variability of spatially averaged JJA precipitation amounts over land in West Africa. The GPCC anomalies illustrate the observed transition between a “wet” period before 1970 and abnormally dry years in the 70s and 80s. The MME fails to clearly exhibit this downward trend of precipitation anomalies. Correlation between the MME relative anomalies and GPCC relative anomalies is 0.382. The ability of the MME to represent the precipitation anomaly varies during the 1960–2005 period : for instance, summers from 1987 to 1993 are fairly well represented whereas in 1984, the MME is totally off-course. Observed relative anomalies range from -17 to $+22\%$. Looking at each individual member, predicted relative anomalies range from -33 to $+35\%$ (not shown). Anomalies are compensated by averaging all members in the MME ensemble mean. This explains the smaller range for predicted anomalies in figure 3.2.

- *The South African region*

As shown in figure 3.1 (d), DJF precipitation in Africa affects mainly areas south of the Equator, with precipitation means over Madagascar exceeding 10 mm. The southern part of the African continent shows a longitudinal gradient in precipitation rates with little to no precipitation on the Atlantic coast and more than 3 mm per day along the Indian Ocean. This is related to a south-eastwards extension of the ITCZ off the Indian Ocean coast of southern Africa. In order to encompass the entire south African region and Madagascar, the area chosen in this paper ranges from 35° S to 10° S and 5° E to 50° E.

Southern Africa precipitation values and spatial distribution strongly vary from one year to the next, partly due to high variability of SSTs in the surrounding oceans. Previous studies

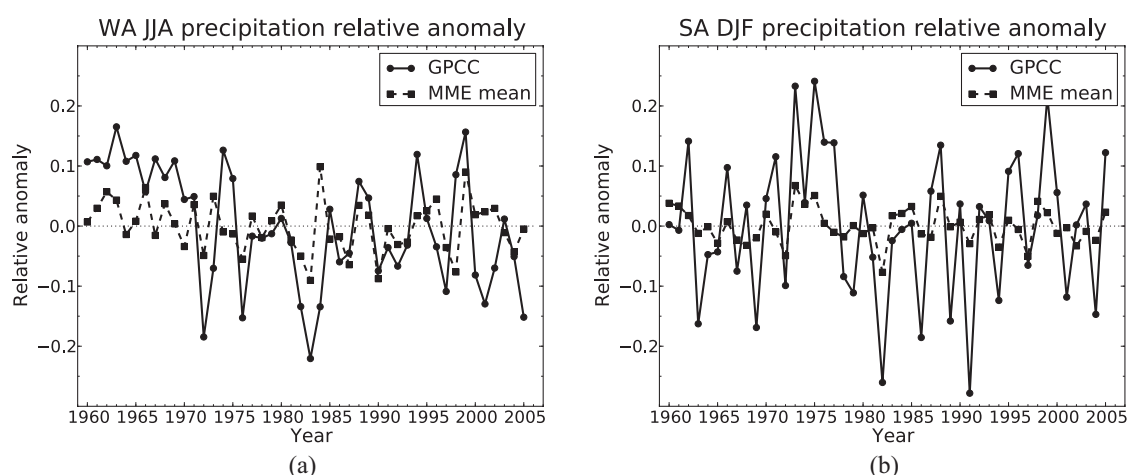


FIGURE 3.2 – Seasonal mean precipitation relative anomaly with reference to the 1960–2005 period for GPCC data and the ENSEMBLES project MME ensemble mean, for spatially averaged West Africa JJA precipitation (a) and southern Africa DJF precipitation (b)

have also shown strong links between ENSO and drought in this region (Rouault and Richard, 2005).

Inter-annual variability of spatially averaged DJF precipitation over southern Africa for GPCC data and modelled by the ENSEMBLES MME is shown in figure 3.2 (b). Correlation between both relative anomaly curves equals 0.624 and is notably higher than for boreal summer in West Africa. The MME duplicates well year-to-year sign and variation of the relative anomaly. Once again, although individual member relative anomalies span the -27% to $+25\%$ range (to be compared with the observed relative anomaly variation between -28% and $+25\%$), MME relative anomalies vary at most from -8% to 7% .

◦ *Greater Horn of Africa long rains and short rains seasons*

Figure 3.1 shows that the structure and intensity of precipitation over the Horn of Africa region varies greatly during the year due to a shift of the ITCZ between the northern and southern hemispheres. Spring (MAM) and fall (SON) rainy seasons appear to be transition seasons between the West African monsoon regime in boreal summer and the South African rains in austral summer. During both of the aforementioned summer events precipitation rates over the Horn of Africa north of the Equator are close to zero. In MAM and SON mean precipitation rates remain small and barely reach 2 to 3 mm per day. MAM is the “long rains” season and shows a lower interannual variability than the SON “short rains” which are strongly influenced by ENSO and the Indian Ocean dipole (IOD). Both seasons have been thoroughly documented in previous papers (see for instance Bowden and Semazzi, 2007 ; Camberlin and Philippon, 2002).

The region chosen in this study, hereafter referred to as “GHA” is a Greater Horn of Africa area including regions affected by both rainy seasons, with latitudes ranging from 13° S to 15° N and longitudes from 25° to 55° E.

The GPCC observed interannual variability of spatially averaged MAM and SON precipitation rates is consistent with the characteristics stated above (see figure 3.3). MAM pre-

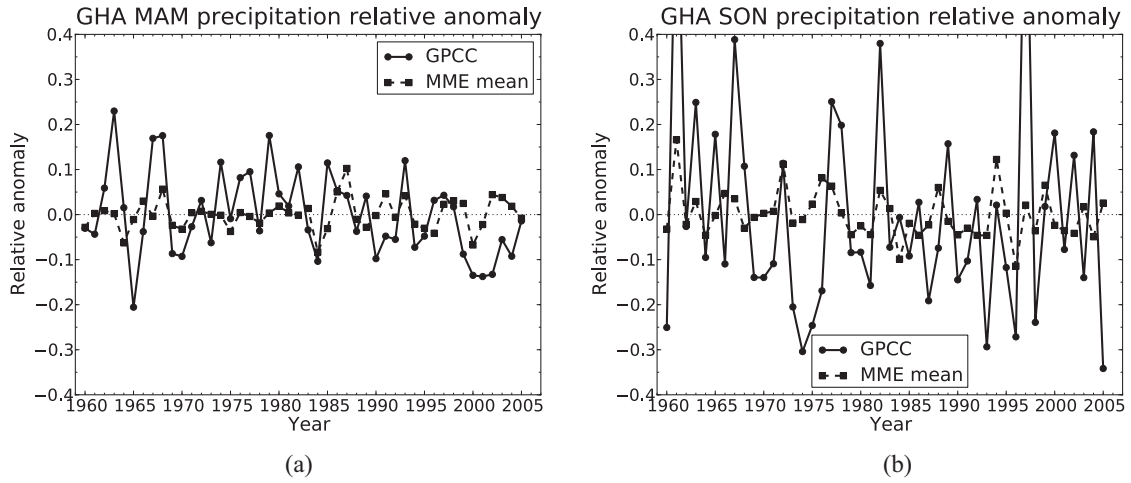


FIGURE 3.3 – Seasonal mean precipitation relative anomaly with reference to the 1960–2005 period for GPCC data and the ENSEMBLES project MME ensemble mean, for spatially averaged Greater Horn of Africa MAM (a) and SON (b) precipitation

precipitation anomalies are smaller than for the SON season, during which positive anomalies reach over 30% four times during the 1960–2005 period. Three of these strong positive anomalies in precipitation could be due to the 1961, 1982 and 1997 positive ENSO-IOD events described in Bowden and Semazzi (2007). The link between SSTs and SON precipitation is a possible explanation for the higher correlation between observed GPCC precipitation anomalies and those predicted by the ENSEMBLES MME. Correlation is worth approximately 0.424 for SON instead of 0.249 for MAM, illustrating the higher predictability of short versus long rains. For MAM, individual members of each model predict higher anomalies than observed, covering a -40% to $+44\%$ range which is twice as large as the observed range. As for SON anomalies, individual members generally underestimate the positive anomalies, which leads to a covered range of -34% to $+60\%$ instead of the observed -34% to $+86\%$. Again, as shown by figure 3.3, the MME ensemble mean predicts much smaller anomalies due to member error compensation.

Results over the GHA region should be interpreted with caution, since fewer rain-gauge stations were available per grid point for the GPCC reanalysis than for the other two regions discussed in this paper (see for instance figure 8 in Schneider et al., 2008).

3.1.2.4 Methodology

The aim of this paper is to provide a range of scores of the ENSEMBLES MME and individual models, as an assessment of ensemble precipitation forecasting skill over the African continent. Alongside currently-used deterministic scores such as root mean square error or correlation coefficients, probabilistic formulations of predictions will be evaluated.

Deterministic scores Deterministic scores are calculated for the MME ensemble mean and individual models. Aside from the anomaly correlation coefficient (ACC) scores calculated for every grid point and season of the 1960–2005 period, a “mean-ACC” score is calculated over

each sub-region for the seasons of interest according to Déqué and Royer (1992). If o_i represents GPCC observed seasonal precipitation over a certain grid point for year i of the time period and p_i the corresponding model prediction, mACC is formulated as follows :

$$mACC = \frac{\overline{\langle (o_i - \bar{o}) \cdot (p_i - \bar{p}) \rangle}}{\sqrt{\overline{\langle (o_i - \bar{o})^2 \rangle} \cdot \overline{\langle (p_i - \bar{p})^2 \rangle}}} \quad (3.1)$$

\bar{X} and $\langle X \rangle$ are averages of variable X respectively over time and space.

A spread-skill ratio was also examined for bias-corrected individual models and the MME to quantify the increase of spread incurred by the multi-model.

Probabilistic forecast and calibration The basic principle of probabilistic forecasts is to use the spread of the different ensembles members, in addition to their mean, in order to issue a forecast. Events chosen here are precipitation exceeding (or remaining below) a certain quantile of climatological precipitation. Threshold quantiles for a given season are defined using the other years of the 1960–2005 period. The way we define here probabilistic forecasts from ensembles is straightforward : the predicted probability is the fraction of individual forecast outputs (out of m months by n ensemble members) which yield precipitation amounts above (or below) the corresponding threshold.

In order to extract as much information from the different forecasts as possible, a quantile-quantile calibration technique was used to correct each forecast data. The choice of such a technique is based on the hypothesis that models are able to provide information on ranks in precipitation rather than accurate amounts. Outputs for a given season are corrected using information from forecasts and observations of the other years of the 1960–2005 period (leave-but-one method). For each grid point, corrected predictions for a given model, season and year are inferred from the deciles of the forecast and observed monthly precipitation values of every other corresponding season of the time period, using linear interpolation between deciles. An example of this correction is given in figure 3.4, which illustrates the calibration method for JJA 2005 monthly precipitation single-model forecasts over a given grid point of WA. Note that this technique corrects the climatology of the members, but not their predictability. In particular, if a model underestimates spread between members, this is not corrected here.

Ranked probability skill score One of the most frequently-used probability scores in meteorology and climatology is the Brier Score, which is calculated as follows for a binary event (Brier, 1950) :

$$BS_i = (y_i - o_i)^2 \quad (3.2)$$

where y_i is the probability forecast of event A during season i and o_i is equal to one if A occurs for season i and zero otherwise. Then BS can be averaged over the n seasons of the period studied. As suggested in Murphy (1973) and Wilks (2006), the Brier Score can be written as the sum of three terms by sorting the forecast-observation pairs according to the J possible forecast values $\{y_j, j \in 1, \dots, J\}$ for a single event. If N_j is the number of times the forecast probability equals y_j the following decomposition is obtained :

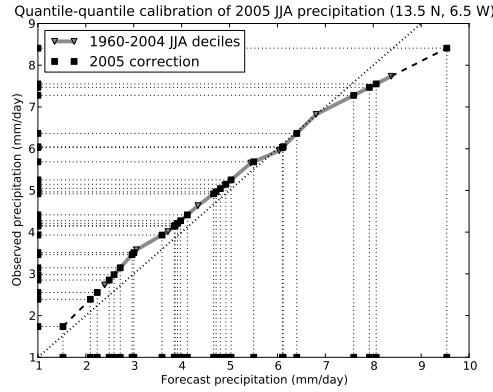


FIGURE 3.4 – Illustration of the quantile-quantile calibration technique used for probabilistic forecasts in the case of JJA 2005 precipitation over grid point (13.5°N, 6.5°W) for an individual model of the ENSEMBLES project. The thick grey curve represents the quantile-quantile plot of forecast (x-axis) and observed (y-axis) JJA monthly precipitation values for the other years of the period (1960 to 2004). Forecast precipitation values for the 9 ensemble members and 3 months of JJA 2005 are placed along the x-axis and corrected values are read along the y-axis (black squares), using the projection (dotted lines) on the quantile-quantile curve and linear interpolation between values (and extrapolation for values beyond quantiles, dashed lines)

$$BS = \frac{1}{n} \sum_{j=1}^J N_j (y_j - \bar{o}_j)^2 - \frac{1}{n} \sum_{j=1}^J N_j (\bar{o}_j - \bar{o})^2 + \bar{o}(1 - \bar{o}) \quad (3.3)$$

$$BS = \text{Reliability} - \text{Resolution} + \text{Uncertainty} \quad (3.4)$$

The reliability term measures a weighted distance between forecast probabilities and observed frequency of the event, giving information on the calibration of the model. It should therefore be as small as possible and can be minimized by forecast postprocessing. Meanwhile, the resolution term is meant to be as large as possible, and estimates the model ability to separate events that have different probabilities of occurrence. A small resolution would mean that for each forecast value y_j the corresponding relative observed frequency of the event \bar{o}_j is close to the overall observed frequency \bar{o} , suggesting that forecast probability values give very little information on the actual outcome of the event. The uncertainty term depends only on observations and their variability : it corresponds to the Brier Score of a forecast based on the climatological distribution of the event. Indeed the reliability and resolution of such a forecast are both worth zero, meaning that climatology forecasts have perfect reliability but no resolution whatsoever.

Toth et al. (2003) show how adding Brier scores for events $A_k = \{X \leq x_k\}$ for K increasing thresholds of a variable X amounts to calculating a discrete ranked probability score, or RPS (Epstein, 1969b). In this paper we chose this method, adding Brier scores for the $K = 9$ precipitation events $A_k = \{\text{seasonal precipitation of year } i \leq x_{k,i}\}$ where $x_{k,i}$ is the k^{th} decile of observations for the other years of the period :

$$RPS = \frac{1}{K} \sum_{k=1}^K BS_k = \frac{1}{K} \sum_{k=1}^K \frac{1}{n} \sum_{i=1}^n (y_{k,i} - o_{k,i})^2 \quad (3.5)$$

Events with 0 and 1 probability are deliberately omitted. RPS ranges from 0 (perfect score) to 1. Calculating this score for the ENSEMBLES models ensemble forecasts (RPS_m) and for the observed climatological distribution over the other years (RPS_c) gives an indication of the relative benefit of forecast use. This relative score is called the RPSS or ranked probability skill score :

$$RPSS = 1 - \frac{RPS_m}{RPS_c} \quad (3.6)$$

RPSS values range from $-\infty$ to 1. Negative values indicate that climatology outperforms the evaluated model in forecasting precipitation quantiles, whereas positive values can be interpreted as the model improvement rate with respect to climatology.

Using the Brier Score decomposition shown in equation (3.4) and the definition of the discrete RPS in equation (3.5) yields a decomposition of the RPS into reliability, resolution and uncertainty terms.

Forecast potential economic value The RPSS compares model forecasts to using the climatological distribution and in doing so estimates the added value of the model. Another approach to probabilistic forecasts evaluation goes even further by computing a model potential “economic value” (EV) in an idealized case study (Palmer, 2002). In this framework, potential economic value for a given event (say, precipitation higher or lower than the k^{th} K -quantile) depends on the hit rate a , false alarm rate b and miss rate c of a model as well as a cost-loss ratio C/L .

In this paper, economic value of a model is defined according to Richardson (2003) as the ratio between savings made by using the actual model instead of climatology and hypothetical possible savings made using a “perfect” model. The mean expense M_m for a model user is given by equation 3.8. The climatology mean expense $M_{\text{climatology}}$ is calculated in leave-but-one mode and the “perfect” model gives an exact probability forecast of the event.

$$EV = \frac{M_m - M_{\text{climatology}}}{M_{\text{perfect}} - M_{\text{climatology}}} \quad (3.7)$$

$$M_m = (a + b) * C + c * L \quad (3.8)$$

For each forecast (real, climatology or “perfect”), the probability for which the user deems preventive action necessary is called probability threshold and is optimized in our study using results from the other years of the time period studied over each grid point. a , b and c therefore depend on the probability threshold and the value of C/L , or cost-loss ratio.

The potential economic value of ENSEMBLES single model forecasts and the MME was evaluated for a set of events and cost-loss ratio (C/L) situations. Events chosen were precipitation exceeding the third quartile of climatology (75th percentile), referred to as E^+ , and precipitation remaining below the first quartile (25th percentile), abbreviated as E^- . We focused on C/L values close to the climatological probability of such events occurring \bar{o} , in both cases worth 0.25. From a user’s point of view, a model forecast is most useful in decision-making for C/L values close to \bar{o} . In other cases, users can rely on their own common sense

to make their mind on whether to spend C on preventive measures or not. Results shown in this paper are overall results for the different regions studied. Rates a , b and c are calculated over all grid points to give an economic value of the models over the region as a whole.

3.1.2.5 Results and discussion

Main results for the ENSEMBLES individual models and MME are presented here, starting with general results over the entire African region and then focusing on the different regions and seasons of interest.

General results over the African continent Figure 3.5 shows seasonal precipitation ACC over the 1960–2005 period for the MME and all four seasons studied in this paper.

In MAM (figure (a)), the ENSEMBLES multi-model failures are located mainly over the Sahara and Central Africa. ACCs over the Greater Horn of Africa region are mainly positive but small. Spatial distribution of ACCs differs between the MAM and JJA seasons (figure (b)). For the latter, scores are higher over the West African region, particularly on the Gulf of Guinea coast. Over the eastern half of the continent ACCs remain positive but weak, and negative scores are found in southern Africa around latitudes from 10 to 20° South where observed mean precipitation is close to zero (see figure 3.1 (b)). In SON (figure (c)), South Africa is the region most affected by negative ACCs. GHA region ACCs are distinctively higher for this season (short rains) than for MAM long rains, most certainly due to the stronger teleconnections with SST indexes discussed earlier. ACCs are higher in South Africa during the DJF season (figure (d)), when the area is affected by the largest precipitation rates (see figure 3.1 (d)). Note however that south of 20° S, correlations are higher where precipitation rates are lower. Positive ACCs are also found over Mauritania and the GHA region for this season.

Generally speaking precipitation ACCs are positive over the regions described earlier during the seasons of interest, even if correlations are weak in the case of MAM precipitation over the GHA region.

Table 3.2 shows the mACC values for individual models and the MME calculated over each region for the season(s) of higher precipitation rates. Calculated mACCs are all positive but reach at most 0.28. Scores for MAM precipitation over the GHA region confirm results shown for ACCs, and are lower than mACCs for the SON season for each individual model and the MME. mACCs obtained for the MME illustrate the interest of combining single model ensembles into a 45-member multi-model ensemble, since scores are higher than single-model scores for all regions and models save for three cases for which a single-model ensemble performs outstandingly well.

Results are also shown for a multi-model 10-member ensemble to test the influence of ensemble size on the MME scores. Each model provides two randomly-drawn members for this reduced multi-model. The score shown in table 3.2 is the mean of 1000 random draws. It gives an indication of score enhancement due to the size increase in the MME. Indeed, over WA the reduced multi-model ranks second when compared to the 9-member models as did the MME, and over GHA and SA it lags behind two individual models with a clearly reduced mACC when compared to the MME. This illustrates that the multi-model performance in precipitation seasonal forecasts resides mainly in the increased number of members of the

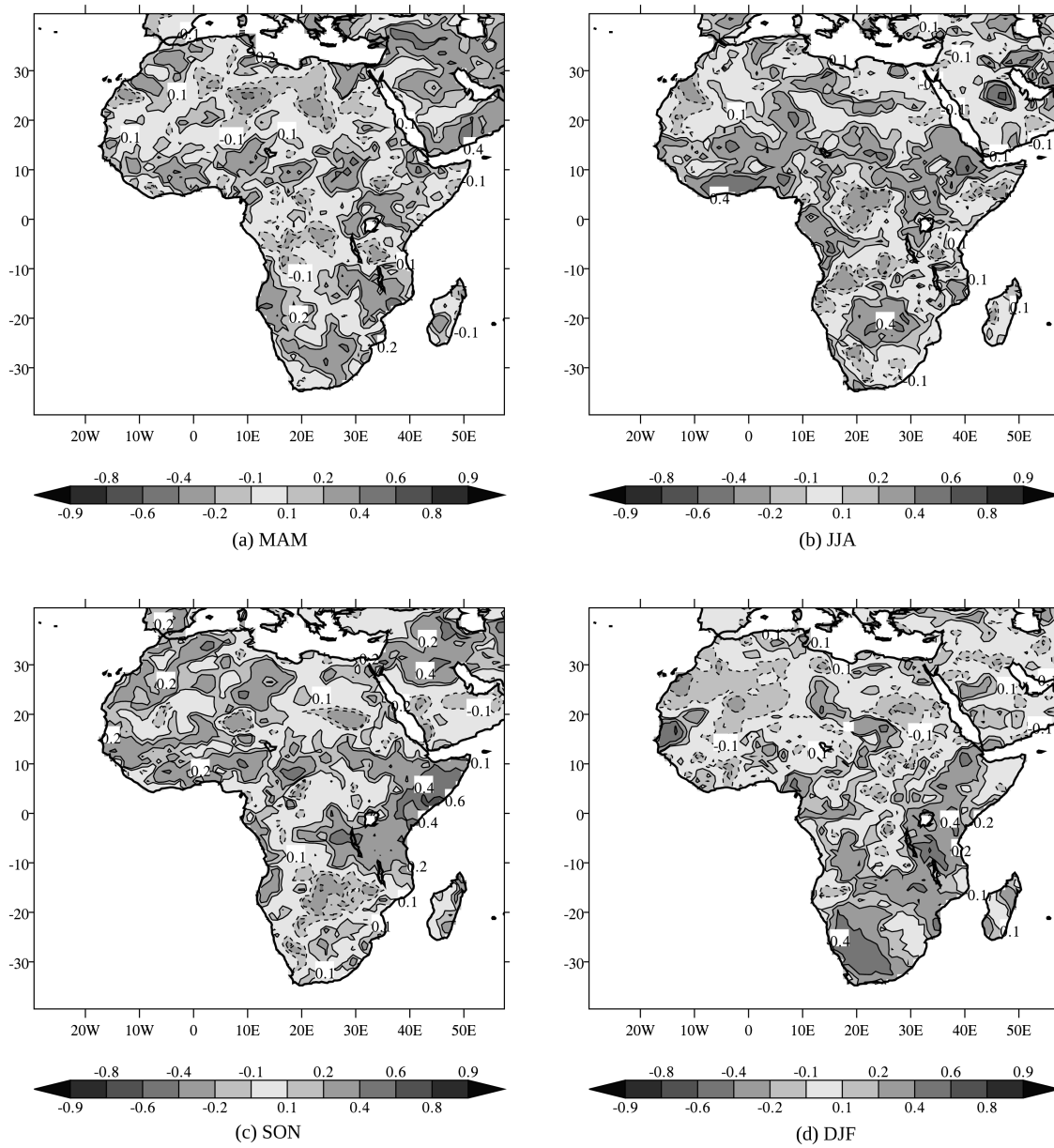


FIGURE 3.5 – Anomaly correlation coefficients over the African continent for the four seasons of interest. Solid and dashed lines indicate positive and negative values, respectively.

TABLEAU 3.2 – mACC values for the regions of study and corresponding seasons of interest over the 1960–2005 period and for three 15 or 16-year sub-periods (*) *see text for more details*

Model	WA JJA	SA DJF	GHA MAM	GHA SON
ECMWF	0.104	0.181	0.030	0.034
IFM-G	0.142	0.138	0.032	0.246
MF	0.130	0.031	0.063	0.285
UKMO	0.227	0.124	0.065	0.101
CMCC	0.149	0.121	0.027	0.056
10-member MM (*)	0.198	0.132	0.057	0.201
MME	0.223	0.189	0.080	0.244
MME 1960–1974	0.300	0.257	0.093	0.353
MME 1975–1989	0.169	0.170	0.097	0.134
MME 1990–2005	0.134	0.190	0.068	0.242

ensemble, especially when one of the single-model ensembles yields poor anomaly correlation coefficients over the region. An apparent drawback in using multi-models is to include poorer models in the ensemble, however a closer look at results over different regions shows that no model is clearly superior to the others in terms of mACC scores, and each model ranks in the top two single models for one region at least.

The last three lines in table 3.2 show MME scores for 15 or 16-year sub-periods of the 1960–2005 time period so as to illustrate a possible variation of skill over time. These scores can only be compared with each other since the length of the time period is in these cases much shorter. They will be discussed in the following paragraphs.

Prediction of West African JJA precipitation More specific results for the West African region are presented here. Figure 3.6 (a) shows a scatter plot of the ensemble spread and ensemble-mean RMSE skill values of each model for every JJA season of the 1960–2005 period. The thin dashed line corresponds to a spread-skill ratio equal to one, the ideal value when ensemble spread is consistent with model error. Individual models have higher errors than their ensemble spread, and unlike RMSE, ensemble spread varies very little from one season to the other. Results are much more satisfactory for the MME (black squares), since for most years ensemble spread is close to RMSE. For some years the MME is slightly over-dispersive, which is never the case for individual models over West Africa. As found in previous studies, multi-model ensembles have a better spread-skill ratio (closer to one) than single model ensembles. However, as stated in Doblas-Reyes et al. (2009) this does not guarantee actual model skill over the region.

On average, the mACC scores shown in table 3.2 are highest for the WA region and range from 0.10 to 0.23, which suggests that a probabilistic formulation can add value to the forecasts by using the entire spectra of information from the different ensemble members. The mACC scores for the sub-periods show a deterioration of mACC for WA JJA precipitation between the 1960–1974 period and the following periods. As explained earlier, the MME fails to capture the downward trend in precipitation anomalies from wet conditions in the 1960s to drought in the 1980s. The 1960–1974 is a relatively wet period during which the MME performs better in terms of mACC.

Figure 3.7 shows the ranked probability skill score for the 45-member MME in its proba-

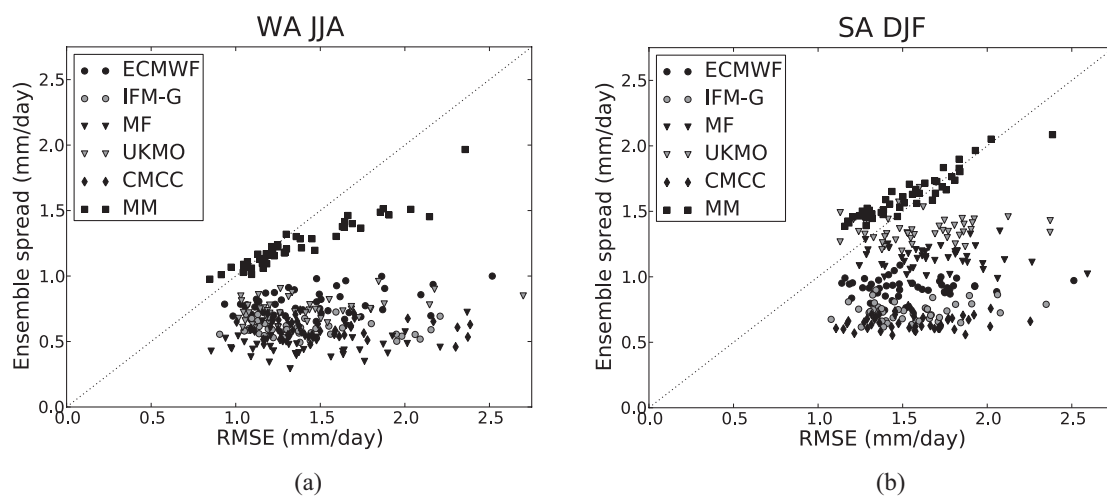


FIGURE 3.6 – Spread-skill scatter plot for the West African region JJA (a) and southern Africa DJF (b) precipitation over the 1960–2005 period. Each point represents spread and ensemble-mean RMSE skill values for a given season and model

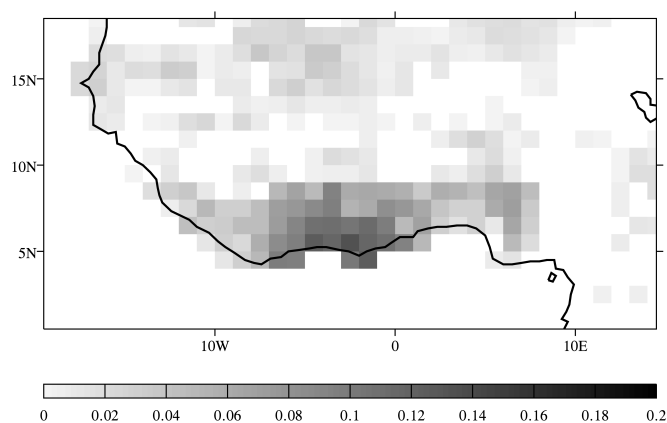


FIGURE 3.7 – RPSS over West Africa for JJA precipitation - only grid points with positive values are shown

bilistic formulation and the A_k events defined earlier. Values are shown at gridpoints where the RPSS is positive, meaning that the MME provides added value with respect to a simple climatological seasonal forecast. As for the ACCs calculated over the region, the best scores are achieved along the northern coast of the Gulf of Guinea. Most grid points of the Sahel region have a positive RPSS although values are very small. Between 9 and 14  N as well as in the eastern part of the area, the MME often ranks below the climatological forecast in terms of RPS. A glance at ACCs calculated for these grid points (see figure 3.5 (b)) shows that they are generally very small and sometimes negative, which illustrates that probabilistic formulation of ensemble forecasts cannot create value over grid points where models have no skill in the deterministic sense. Generally speaking, RPSS values are small and range from -10% to 13.5% with an average over WA of 1.0% .

Table 3.3 shows the RPSS and RPS spatial average and for each single model and the MME as well as the reliability-resolution terms of the RPS discussed earlier. These terms are divided by the RPS uncertainty term. Scores for WA are shown in the upper left corner of the table. The MME is the only model for which the spatial average of RPSS is positive over the West African region, meaning that single models alone are less efficient than climatology in predicting precipitation ranks. When compared to single model RPSS maxima (13.1%) and minima (-62.4%) over the area, the multi-model clearly improves the minimum RPSS over the region whereas the maximum RPSS value is only slightly larger than the highest single model maximum score. This shows that multi-model additional information is particularly useful over grid points where single models experience difficulties and won't further improve scores over grid points where precipitation predictions are already skillful. This conclusion is corroborated by comparing figures of RPSS values over the area for single models and the MME (not shown).

The reliability-resolution terms in table 3.3 show that the multi-model improves the single-model RPSS by dramatically increasing model resolution. This positive effect compensates a smaller increase in the reliability term, meaning that the multi-model calibration is poorer than that of single models.

Economic value of the ENSEMBLES single models and MME for events E^- and E^+ over West Africa during the JJA season for different cost-loss ratio values is shown in figure 3.8. C/L values range between 0.05 and 0.45 with a 0.025 step. For both events, EV is maximum for $C/L = \bar{o}$ and reaches at most 10.4% for E^+ and 6.1% for E^- . In most cases \bar{o} is the only C/L value for which models reduce expenses with respect to climatological forecasts. The three models with highest values are the same for both events, and values are generally higher for E^+ than E^- . Unlike results shown for RPSS, the multi-model does not improve economic value with respect to individual models for both events.

Results over the Greater Horn of Africa Results over GHA vary between the MAM and SON seasons. Mean ACC scores shown earlier suggested that seasonal precipitation predictability was higher for SON than MAM. A comparison of figures 3.9 (a) and (b) does not lead necessarily to the same conclusion : both scatter plots show similar spread-skill ratios for both seasons, with spread and RMSE generally lower during the MAM season. As for WA in JJA, the multi-model helps broaden model ensemble spread while maintaining similar RMSE levels. For both seasons and for most years of the 1960–2005 period the multi-model is slightly over-dispersive.

TABLEAU 3.3 – Single model and MME spatial average RPSS (%), RPS, and reliability - resolution decomposition (values divided by uncertainty) for the different regions and seasons of interest over the 1960–2005 period

Model	WA JJA				SA DJF			
	RPSS (%)	RPS	Rel	Res	RPSS (%)	RPS	Rel	Res
ECMWF	−2.71	.188	.29	.26	−0.45	.184	.27	.27
IFMG	−2.21	.187	.26	.23	−2.12	.187	.27	.25
MF	−5.65	.194	.36	.30	−6.00	.194	.34	.28
UKMO	−0.38	.184	.28	.28	−2.18	.187	.27	.25
CMCC-INGV	−8.32	.199	.39	.30	−2.25	.187	.29	.26
MME	0.98	.182	.41	.42	1.44	.180	.41	.42

Model	GHA MAM				GHA SON			
	RPSS (%)	RPS	Rel	Res	RPSS (%)	RPS	Rel	Res
ECMWF	−4.16	.191	.36	.31	−8.08	.198	.41	.33
IFMG	−3.96	.191	.36	.32	−4.59	.192	.37	.33
MF	−4.70	.192	.37	.32	−3.95	.191	.39	.35
UKMO	−3.53	.190	.35	.32	−6.78	.196	.41	.34
CMCC-INGV	−4.48	.192	.37	.32	−6.97	.196	.39	.32
MME	−1.25	.186	.44	.43	−2.86	.189	.49	.46

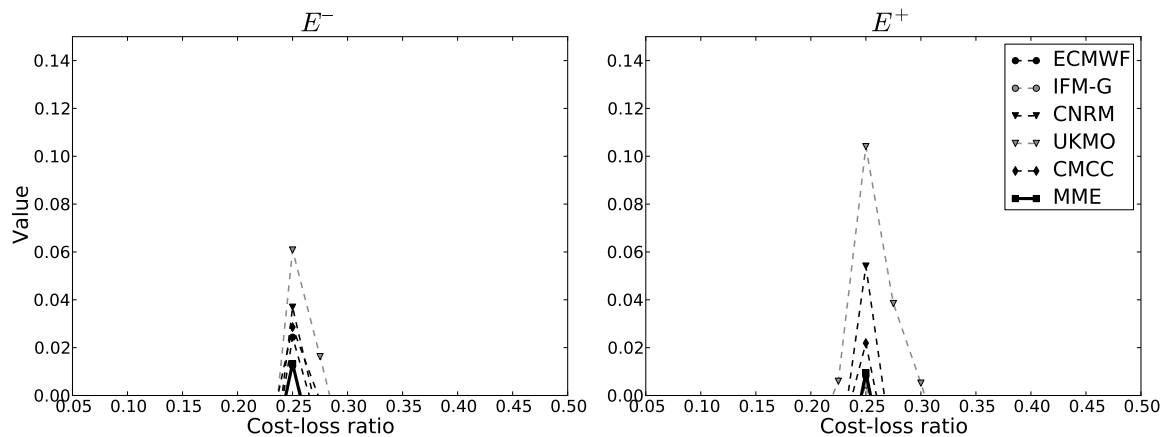


FIGURE 3.8 – Economic value of ENSEMBLES project single models and MME (thick line) for events E^- (left) and E^+ (right) and different cost-loss ratio values (see text for more details) for predictions of JJA rainfall over West Africa

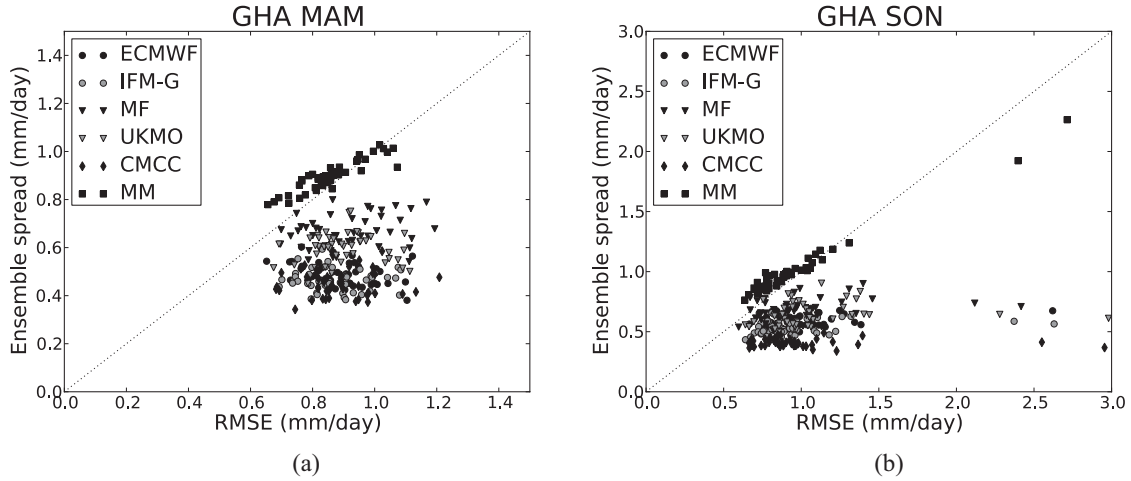


FIGURE 3.9 – Same as figure 3.6 but for Greater Horn of Africa MAM (a) and SON (b) precipitation. Note that axis scales are not the same for both figures

The evolution of mACC for three sub-periods shown in table 3.2 differs between both seasons. In MAM, no striking difference is found between the sub-periods, although 1990–2005 scores are slightly lower : the three sub-periods show low anomaly correlation coefficients for the MME over the GHA region. In SON, mACCs are notably higher for the 1960–1974 sub-period (0.353), when mean precipitations are higher than average, and lower for the 1990–2005 sub-period (0.134), when mean precipitations are slightly lower, although strong interannual variations in precipitation anomalies make further interpretation of these results quite difficult (see figure 3.3 (b)).

RPSS calculations over GHA are shown in figures 3.10 (a) and (b) for the MAM and SON seasons respectively. An immediate conclusion when looking at both figures is that the MME precipitation deciles predictions are much better for SON than for MAM. There are many more grid points with positive RPSS in the case of SON (69% versus 31.9% for MAM). The spatial distribution of positive RPSS is once again consistent with results for anomaly correlation coefficients. However, RPSS values shown in table 3.3 (lower half) show that average scores are lower for SON than MAM. This is due to highly negative RPSS scores over the northwest corner of the GHA domain. All models tend to always predict positive precipitation rates over these gridpoints where little to no precipitation is actually observed. In spite of the quantile-quantile correction each model exhibits very poor skill over this sub-region, penalizing the overall score over GHA for SON.

Individual models and MME economic values for events E^- and E^+ were calculated over GHA for both seasons. Results for E^+ during SON are shown in figure 3.11. Each model has a positive EV for $C/L = \bar{o}$, and values reach over 15% for one single model. The multi-model ranks second among all models, contrasting with results over WA.

For E^- in SON only one model has a positive EV (of only 3%) over GHA. Similarly disappointing results are found over GHA during MAM where very few models show hints of economic value and multi-model values remain negative for all cost-loss ratios. These results are not surprising given the poor performance of probabilistic and deterministic forecasts over the region, and are therefore not shown here.

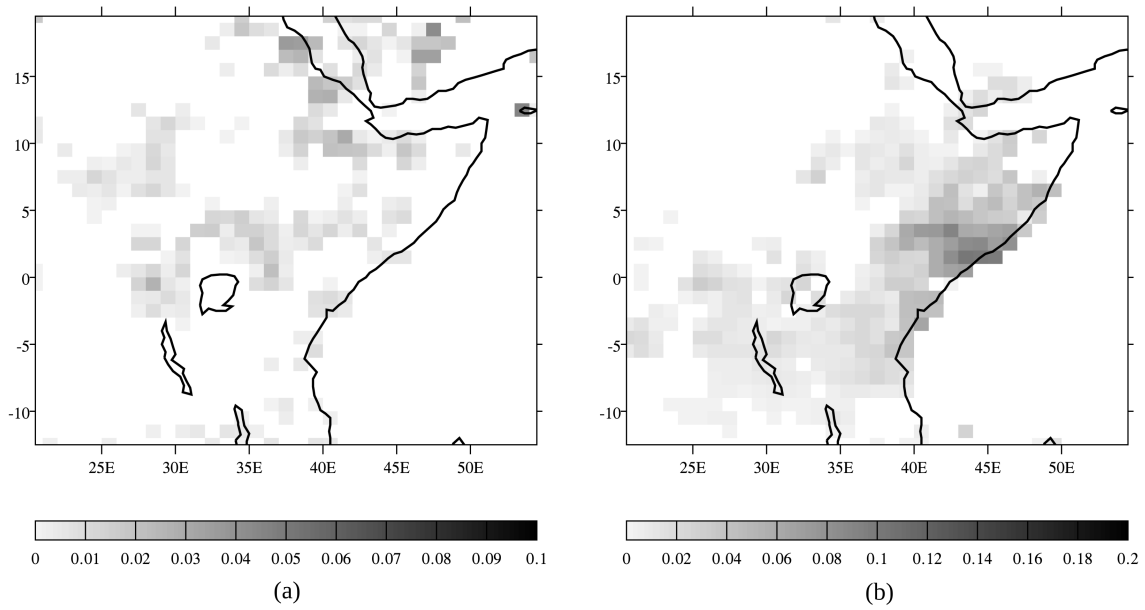


FIGURE 3.10 – Same as figure 3.7 but for Greater Horn of Africa MAM (a) and SON (b) precipitation

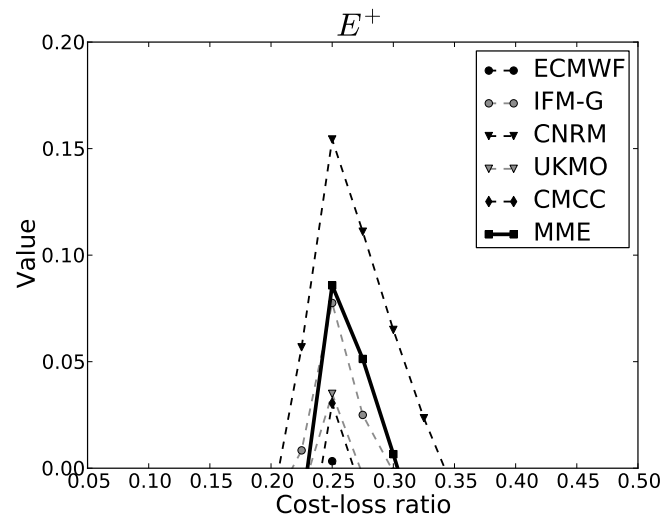


FIGURE 3.11 – Economic value of ENSEMBLES stream 2 single models and MME (thick line) for event E^+ and different cost-loss ratio values (see text for more details) for predictions of SON rainfall over GH

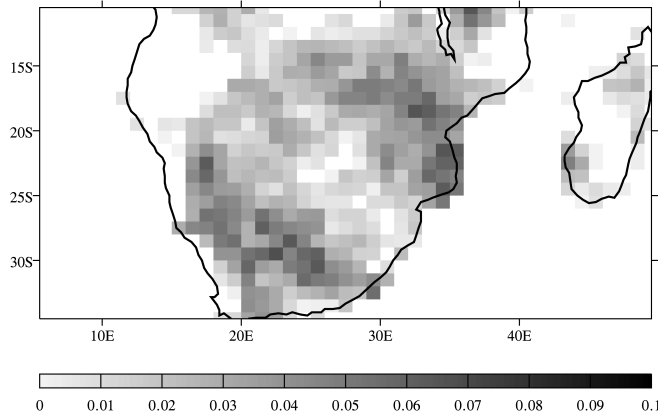


FIGURE 3.12 – Same as figure 3.7 but for southern Africa DJF precipitation

South African DJF precipitation forecast skill A spread-skill scatter-plot for DJF precipitation over the SA region is shown in figure 3.6 (b). Results are quite similar to those observed over WA in boreal summer, even if ensemble spread and RMSE values are generally higher. Ensemble spread interannual variations for individual models are small, leading to a mainly horizontal dispersion of points in the scatter plot. This is corrected by using the MME for which the alignment of points is closer to the idealistic diagonal line. As for GHA, switching from single models to MME leads to a better spread-skill ratio and slightly over-dispersive ensembles.

When compared to WA JJA, mACC values for SA DJF (table 3.2) are poorer for 4 out of 5 single models and the multi-model, yet the MME mACC value (0.189) gives hope that a probabilistic formulation may yield positive results over the region. It is worth noticing that SA DJF ACC values are less noisy in space than WA JJA (see figures 3.5 (b) and (d)) although spatial distribution of precipitation amounts is heterogeneous in both cases. Table 3.2 shows that mACC over SA is higher for 1960–1974 than for the other two sub-periods examined, as for WA JJA and GHA SON. However, unlike WA JJA no clear trend in precipitation can be seen in observations or in the MME predictions (see figure 3.2(b)).

Figure 3.12 shows the ranked probability skill scores of the MME over each grid point for SA DJF precipitation. Results averaged over the area are shown in the upper right-hand corner of table 3.3. The MME RPSS averaged over land gridpoints is positive, and even slightly larger than for WA JJA, although maximum and minimum values over the area are quite smaller than those calculated for West Africa (not shown). More strikingly, positive RPSS values are found over approximately 75% of the SA region (for WA JJA, RPSS is positive for only 57% of the grid points).

Results for model EV over southern Africa for dry (E^-) and wet (E^+) events are given in figure 3.13. Unlike over WA, the multi-model improves economic value of most single models for E^- , with maximum value reaching 7.3% of a perfect model score. Values for E^- are slightly higher than over WA, and the best models are not the same. For E^+ results are quantitatively similar for single models, yet the multi-model EV is 50% lower than that of the highest ranking model. Results for the multi-model are still much better than over West Africa.

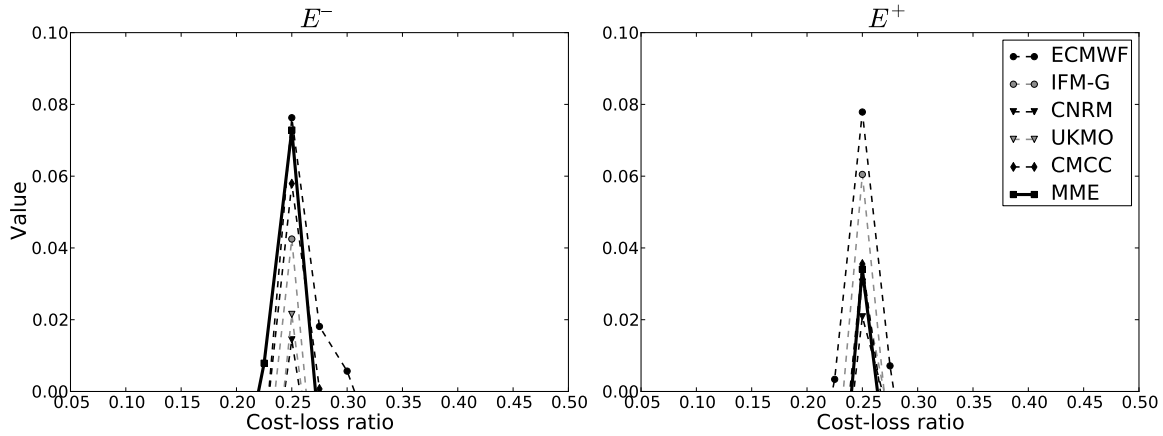


FIGURE 3.13 – Same as figure 3.8 but for predictions of DJF precipitation over SA

Links between SST and precipitation Results shown in the previous paragraphs are generally better for regions and seasons where precipitation is strongly linked to oceanic surface conditions, as illustrated for instance by differences in mACC between MAM and SON precipitation forecasts over the GHA region.

To further investigate links between sea-surface temperature and precipitation prediction by the MME, several additional score calculations and model-observation comparisons were led. The ERA-40 prescribed SST data was used for years from 1960 to 1988 included, and ERA-Interim prescribed SST data covered the remaining 1989–2005 period. Note that ERA-40 and ERA-Interim SST data are almost the same from 1989 to 2001. Prescribed sea-surface temperatures are based on the HADISST1 dataset prior to November 1981 and NOAA/NCEP SST products thereafter (Uppala et al., 2005).

◦ *Covariance between precipitation anomalies and simultaneous SSTs*

As in Guérémy et al. (2005), covariances between seasonal SSTs and standardized anomalies of precipitation over the regions studied were calculated to exhibit the SST anomalies related to precipitation anomalies and therefore possible forcing regions. Results are shown in figure 3.14, and can indicate the ability of the model to replicate observed links between seasonal precipitation anomalies over the GHA, WA and SA regions and simultaneous sea-surface temperatures around the globe.

Covariances between MAM SSTs and MAM GHA precipitation anomalies show no distinct patterns in the observation datasets (figure 3.14 (a)). Camberlin and Philippon (2002) suggested that MAM GHA precipitation links with SSTs over the Tropical Pacific were difficult to assess at a seasonal scale since phase shifts in ENSO often occur during this season. The MME covariances shown in figure 3.14 (b) are positive over the Eastern Tropical Pacific, meaning that higher precipitation over GHA is correlated with higher SSTs over this region in the model. Covariances over the Atlantic and Indian Oceans are also inconsistent with observations, illustrating possible sources of model error.

Results for other regions and seasons are generally better. Figures 3.14 (c) and (d) show that the MME finds the same sign of covariances as the observations over the tropical Pacific and the Gulf of Guinea, although values differ. As in the observed data, higher-than-average SSTs over the Gulf of Guinea and lower-than-average SSTs over the Eastern Tropical Pacific

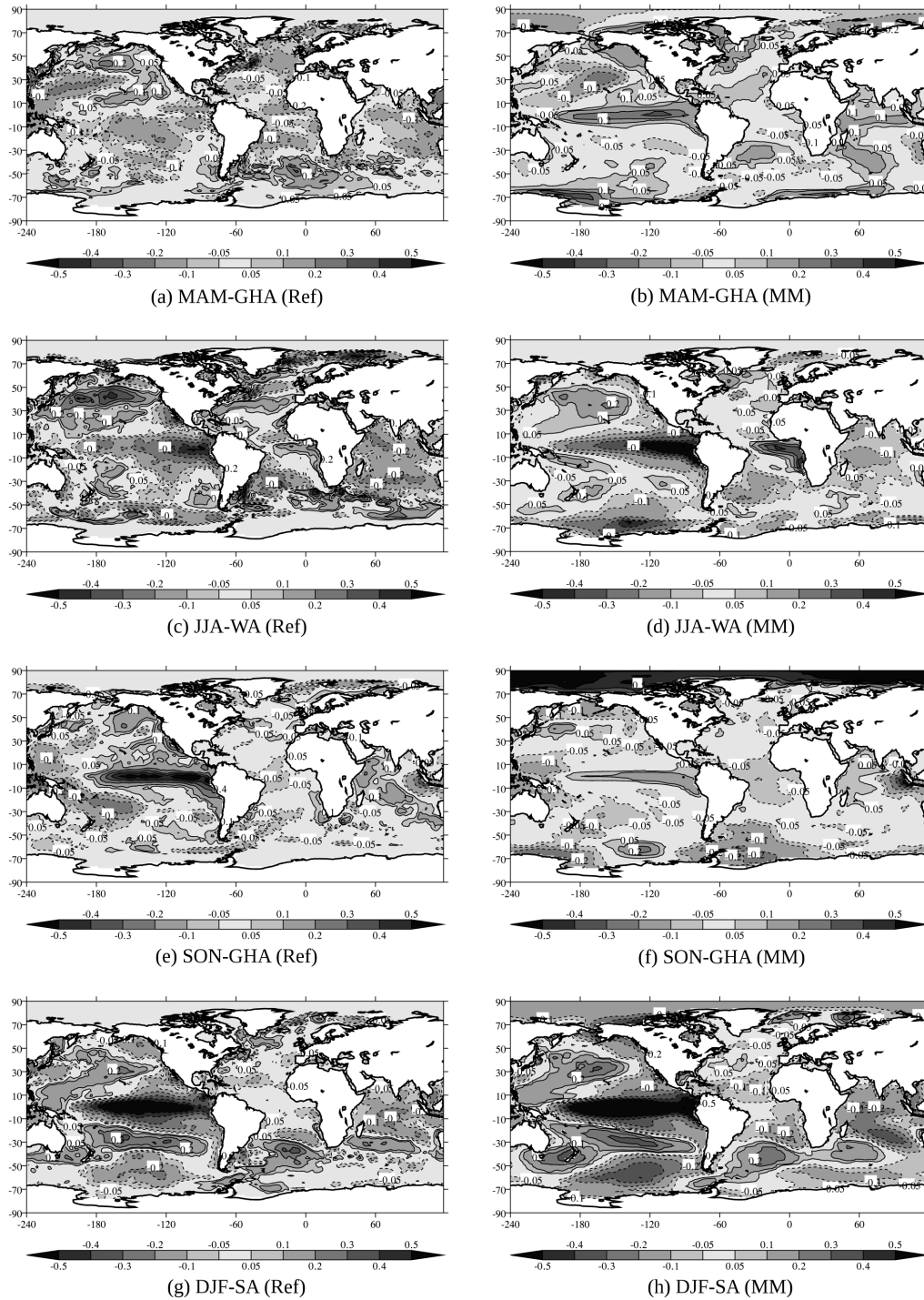


FIGURE 3.14 – Covariance between SSTs and standardized precipitation anomalies over GHA (for MAM, SON) WA (for JJA) et SA (for DJF). The left column shows covariances between ERA SSTs and GPCC precipitation and the right column shows the MME covariances. Values are in Kelvin, solid and dashed lines indicate positive and negative values, respectively.

TABLEAU 3.4 – Classification of DJF seasons in warm, average and cold years categories for Niño 3.4 SST. (The year corresponds to the month of December)

Warm years	1963 1965 1968 1969 1972 1976 1977 1979 1982 1986 1987 1991 1994 1997 2002 2004
Average years	1959 1960 1961 1962 1966 1974 1978 1980 1981 1989 1990 1992 1993 1996 2001 2003
Cold years	1964 1967 1970 1971 1973 1975 1983 1984 1985 1988 1995 1998 1999 2000 2005

TABLEAU 3.5 – mACC values for the regions of study and corresponding seasons of interest over the 1960–2005 period for the MME over warm, average and cold DJF SST years

ENSO category	WA JJA	SA DJF	GHA MAM	GHA SON
Warm years	0.214	0.174	0.068	0.108
Average years	0.229	0.047	−0.050	0.381
Cold years	0.201	0.151	0.190	0.155

lead to positive anomalies in precipitation over the WA box. SON SST covariance with GHA precipitation anomalies shows a strong positive pattern over the Eastern Tropical Pacific (figure (e)) which also appears for the MME but with a smaller amplitude (figure (f)). Negative covariances just west of Indonesia are well replicated. Out of all seasons and regions, the one for which the MME depicts observed covariances best is DJF SA. A Pacific Ocean SST latitudinal quadrupole appears in both figures (g) and (h), and the MME shows similar structures for South Atlantic and Indian Ocean covariances.

These results can be linked to the MME performances observed earlier. GHA in MAM is the region for which scores are lowest, and the MME seems to yield unobserved patterns between SSTs and precipitation anomalies. On the other hand, for DJF SA, the MME depicts well both patterns and values of covariances, and this is the region and season for which the best probabilistic scores are found.

◦ Dependency of results on ENSO

Previous studies have shown that seasonal prediction skill is closely related to ENSO amplitude (Wang et al., 2009). In order to study the dependency of the earlier presented results with ENSO, the 46 years of the time period were split into three sets depending on the tercile of DJF Niño 3.4 mean observed SST in the current year. DJF 1959–1960 was also included in the time period to study 1960 seasonal forecast skill.

The three sets obtained are listed in table 3.4. These sets are the same if the Niño 3 region is used instead of Niño 3.4, and are quite similar to recorded El Niño and La Niña episodes. The choice of DJF terciles as a criterion instead of usual ENSO indices is dictated by the need to obtain three sets of same size, so scores such as ACC are comparable.

mACC results were calculated for each region and season for these three categories, using for WA JJA and GHA MAM the preceding DJF season, and for GHA SON and SA DJF the following or synchronous DJF season as a classification criterion (for instance, for DJF 1997–1998, SON 1997, DJF 1997–1998, MAM 1998 and JJA 1998 scores are related to the “warm years” category). Results for these mACC are presented in table 3.5. For WA JJA, results are similar for all three categories, indicating that the same level of correlation is maintained in

non-ENSO years. This tends to illustrate the predominance of local and synchronous ocean conditions such as JJA Gulf of Guinea SSTs over large-scale teleconnections in MME forecast skill over the WA region as a whole.

Results over GHA and SA are more contrasted. MME mACCs are higher over the average years sub-period than for warm and cold years for GHA SON precipitation prediction, but the opposite results are found for GHA MAM and SA DJF. Figures 3.14 (g) and (h) showed that anomalies in DJF precipitation over SA were negatively correlated to DJF SSTs over the tropical Pacific, and that the MME exhibited a covariance pattern similar to observations. The improved mACC results obtained for SA DJF during warm and cold years are consistent with these former results.

Results for MAM mACCs are somewhat surprising since no clear connection between ENSO and seasonal-scale precipitation over the entire GHA domain has been documented. Grid-point temporal ACCs for the three sub-categories of years over GHA reveal that strong positive ACCs are found during cold years over equatorial East Africa and during warm years over the northwestern corner of the GHA domain and north of Kenya and Uganda (not shown). For average years, a negative mACC is found, consistent with scores found over most regions of the GHA domain. For SON, high ACCs (over 0.7) for average years are found over Somalia and along the coast, and positive ACCs mainly south of 4  N for inland areas (not shown). During warm and cold years, ACC patterns differ, with positive values located mainly north of the Equator for cold years and over the Horn of Africa and south of the Equator (but with smaller values) for warm years. Lower ACCs for warm and cold years in SON could be explained by the fact that the MME forecasts are initialized in August, and early signs of warmer or colder than average DJF SST conditions may not be well captured by the MME. It is interesting to note that in the case of warm years (generally giving rise to higher SON precipitation rates over GHA), lower mACCs for the MME ensemble mean do not translate into absence of skill of the probabilistic MME in forecasting wet events (such as E^+ studied for economic value scores).

3.1.2.6 Summary and conclusions

Seasonal forecast skill is often assessed over tropical regions for surface temperature variables. This paper gives an overview of the ENSEMBLES stream 2 seasonal forecast project : single models and multi-model ensemble seasonal precipitation forecast skill is assessed over the African continent. The aim of this study is to illustrate potential predictability of seasonal precipitation using only simple and robust model calibrations.

Results over specific regions during higher-than-average precipitation or monsoon seasons are quite contrasted. For all regions in this study the 45 member MME considerably increased ensemble spread without deteriorating model errors. Over West Africa during JJA, results are encouraging. The ENSEMBLES MME manages to grasp some interannual variability and has added value with respect to climatology in terms of RPS not only over the Gulf of Guinea but also over the Sahel (although with much lower scores), which is not the case for most individual models. In southern Africa for the DJF season, positive RPS skill is found over most grid points of the region. Scores are much more disappointing over the Greater Horn of Africa. For the SON “short rains” season, more grid points with positive RPSS are found than for the MAM “long rains” season. However, average scores for SON are lower due to strongly negative RPSS over the northeastern part of the GHA domain. Generally speaking, results are

thus better for regions and seasons where precipitation is strongly linked with oceanic surface conditions.

In terms of EV, positive values were found for both single models and the MME over the West African and southern African regions for cost-loss ratios equal to the probability of occurrence of the precipitation events studied. This illustrates that even for a variable highly dependent on parameterizations such as precipitation, GCMs can show additional skill when compared to climatology. These results are encouraging and contrast with the reputation GCMs have been given over the years when it comes to precipitation forecasts outside the Tropical Pacific area. They also suggest room for improvement using adaptation techniques such as model-output statistics.

Maps of MME global SST covariance with precipitation anomalies over the regions studied were compared to observations, and showed that the ENSEMBLES MME generally replicated the main SST patterns linked with simultaneous precipitation anomalies, save for the GHA region in MAM where no clear observed patterns were found. Deterministic scores calculated for warm, average and cold years of DJF SSTs over the Niño 3.4 region illustrate possible links between ENSO and MME skill, especially over the SA region where mACCs were higher for warm and cold years than for average years. No clear difference in skill was found between the three sets of years for the WA box in JJA. This could be due to the fact that the WA box studied in this paper comprises both the Gulf of Guinea, strongly influenced by local SSTs, and the Sahel for which precipitation is known to be linked to ENSO.

Probabilistic evaluations of seasonal precipitation predictability are more user-related than general deterministic scores such as anomaly correlation. It is worth noticing however that positive results for probabilistic forecasts were found only over regions where deterministic skill had been demonstrated, suggesting that the use of information from all of the ensemble members is fruitful only where the ensemble mean forecast shows hints of performance. Results might depend on the event studied and the number of quantiles chosen for calibration, but additional calculations using terciles instead of quartiles for EV showed similar results. The fairly long time period during which data is evaluated (45 years) argues for robustness of results, however it is worth noting that most scores are low (especially the RPSS) compared to SST prediction in the tropical Pacific. The usefulness of MME seasonal forecasts must therefore be qualified in application studies implying for instance water management or health-related decision-making.

ACKNOWLEDGMENTS

The ENSEMBLES data used in this work was produced by the European Commission FP6 Integrated Project ENSEMBLES (Contract number 505539). The Global Precipitation Climatology Centre (GPCC) is operated by Deutscher Wetterdienst (DWD) under the auspices of the World Meteorological Organization (WMO). ERA-40 and ERA-Interim SST data were supplied by ECMWF. We are grateful to two anonymous reviewers who contributed to improve significantly the manuscript.

Modèle	mACC		RPSS	
	Guinée	Sahel	Guinée	Sahel
ECMWF	0.167	0.078	-0.030	-0.029
IFMG	0.175	0.182	-0.020	-0.006
MF	0.209	0.073	-0.044	-0.054
UKMO	0.290	0.205	0.006	-0.010
CMCC-INGV	0.278	-0.002	-0.072	-0.085
MM	0.322	0.171	0.035	0.002

TABLEAU 3.6 – Scores mACC et RPSS sur les zones Guinée et Sahel des modèles individuels et du multi-modèle ENSEMBLES en JJA.

3.2 Le cas de l’Afrique de l’ouest : prévisibilité sur le Golfe de Guinée et le Sahel, et liens avec les TSO

Afin de compléter l’étude des prévisions de précipitations sur la région de l’Afrique de l’ouest présentée dans l’article, la région a été séparée en deux domaines distincts. Le premier domaine se situe au sud de 10°N et englobe les régions voisines du Golfe de Guinée principalement influencées par l’océan Atlantique équatorial ; l’autre, au nord de 10°N, comprend les zones du Sahel et le sud du Sahara dont les précipitations, moins prévisibles et fortement variables à l’échelle décennale, ont une variabilité interannuelle fortement dépendante de l’ENSO (Giannini et al., 2003).

3.2.1 Le Golfe de Guinée et le Sahel : deux régions contrastées

La figure 3.15 reprend les résultats montrés dans le papier publié dans Tellus A pour le RPSS du multi-modèle ENSEMBLES, mais calculé sur les terciles de précipitations. Elle illustre très nettement l’écart entre les deux régions des performances des prévisions de précipitations. Ces écarts sont quantifiés dans le tableau 3.6 où les mêmes scores de mACC et de RPSS sont calculés pour les modèles individuels et le multi-modèle ENSEMBLES. Mis à part le modèle de l’IFM-Geomar, les scores sont meilleurs sur la partie Guinée que sur le Sahel. Cela se retrouve également dans les scores du multi-modèle.

L’étude de Philippon et al. (2010) portant sur les prévisions des modèles du « stream 1 » d’ENSEMBLES (1990–2001) montre également les différences de performances entre la prévision des pluies sur le Sahel et celle sur le Golfe de Guinée. Dans leur étude, les auteurs ont procédé différemment. Plutôt que de calculer les scores sur des boîtes comme celles définies ici, ils évaluent les corrélations entre les séries temporelles des valeurs propres des deux premières EOF du champ des précipitations des modèles et celles des observations. Cette méthode permet de moins pénaliser les modèles qui situent les pluies sahéliennes trop au sud, ce qui est le cas par exemple des modèles du CEPMMT et de Météo-France utilisés dans ENSEMBLES. Malgré cette décomposition en composantes principales, les prévisions restent moins bonnes sur le Sahel que sur la Guinée, que ce soit en termes de scores déterministes ou probabilistes.

Contrairement aux pluies guinéennes, régies essentiellement par les anomalies de TSO sur l’Atlantique tropical et le Golfe de Guinée, les pluies sahéliennes sont également soumises à l’influence de l’ENSO. Dans la partie suivante, nous détaillons les liens entre précipitations sur

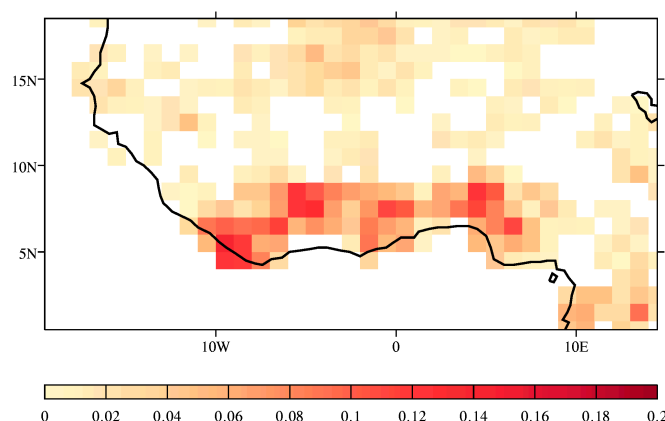


FIGURE 3.15 – RPSS du multi-modèle ENSEMBLES calculé sur les terciles des précipitations sur la période JJA 1960–2005.

ces zones et les TSO, et leur représentation dans les modèles du projet ENSEMBLES.

3.2.2 Liens entre précipitations sur l'Afrique de l'ouest et TSO, et représentation dans les modèles

On procède comme dans l'article au calcul des scores de mACC en séparant les saisons JJA 1960–2005 en trois catégories en fonction du tercile de TSO sur la région Niño 3.4 en DJF. Contrairement à ce qui est fait dans l'article, on utilise les terciles de TSO de la saison DJF suivant la saison de mousson. En effet, les travaux de Joly et Voldoire (2009) montrent que dans les réanalyses, les téléconnexions entre l'ENSO et les précipitations sur le Sahel ont lieu lorsque la saison d'été coïncide avec le début d'un événement Niño ou Niña, ou un événement Niña persistant, mais très rarement avec des fins de Niño et Niña. Les scores de mACC présentés ici sont des scores relatifs, dans la mesure où les anomalies observées et prévues sont calculées par rapport à l'ensemble de la période de référence. Les résultats sont montrés pour les modèles individuels et le multi-modèle dans le tableau 3.7. Sur le Sahel, la plupart des modèles ont des scores contrastés selon la catégorie de TSO. De manière générale, c'est pour la catégorie de TSO la plus chaude que les scores sont les meilleurs, et on retrouve cette tendance dans les scores du multi-modèle. Pour la boîte Guinéenne, le contraste est moins marqué, même si les scores sont généralement meilleurs pour la catégorie de TSO médiane (Neutre), ce qui se traduit par une mACC du multi-modèle plus élevée dans ce cas.

Ces résultats sont à interpréter avec précaution, ce pour plusieurs raisons. Pour commencer, on a choisi d'utiliser les terciles de TSO sur la boîte Niño 3.4 afin de définir les trois catégories appelées abusivement « Niño », « Niña » et « Neutre ». Si ce choix permet de calculer des scores sur des sous-échantillons d'années de même taille, il a le défaut d'atténuer l'impact éventuel d'un fort signal Niño sur la prévisibilité, en incluant les années de fort ENSO parmi d'autres années où l'anomalie de TSO est nettement moins marquée. Des scores si faibles, qui plus est calculés sur des périodes d'au plus 13 ans, ne sont pas significatifs. Par ailleurs, l'étude de Joly et Voldoire (2009), portant sur les modèles de climat couplés utilisés pour l'intercomparaison CMIP3, a montré que plusieurs modèles (dont le modèle CNRM-CM3) représentaient une

3.2 Prévisibilité sur le Golfe de Guinée et le Sahel, et liens avec les TSO

Modèle	Sahel			Guinée		
	Niño	Neutre	Niña	Niño	Neutre	Niña
ECMWF	0.15	0.03	0.05	0.11	0.18	0.21
IFMG	0.16	0.17	0.22	0.17	0.10	0.24
MF	0.07	0.04	0.10	0.19	0.22	0.22
UKMO	0.34	0.18	0.13	0.35	0.27	0.24
CMCC-INGV	0.18	0.00	-0.14	0.18	0.37	0.30
MM	0.30	0.13	0.10	0.29	0.38	0.32

TABEAU 3.7 – Scores mACC en fonction des terciles de TSO sur la région Niño 3.4 pendant la saison DJF qui suit la saison JJA étudiée, sur le Sahel et la Guinée.

téléconnexion entre les précipitations sahéliennes et la fin d’un événement ENSO plutôt que le début. Les liens entre précipitations sahéliennes et l’ENSO ne sont donc pas correctement représentés dans ces modèles. S’ajoute à cet écueil le fait qu’en mode prévision saisonnière, le modèle couplé doit prévoir l’arrivée prochaine d’un événement de type Niño ou Niña. Il y a donc encore de nombreux progrès à faire dans le but de prévoir correctement les précipitations sur le Sahel à l’échelle saisonnière.

Les scores obtenus par le multi-modèle les années de type « Niño » sur le Sahel ainsi que la figure 3.14 de l’article publié dans *Tellus A* montrent toutefois qu’il y a lieu de croire que certains liens entre TSO et précipitations sont correctement représentés dans celui-ci. La figure 3.16 montre les covariances (en degrés) entre TSO et anomalies de précipitations en JJA sur les deux zones de l’Afrique de l’ouest. Dans les données de référence, les anomalies positives de précipitations sur le Sahel sont liées à des TSO froides sur le Pacifique Est Tropical, et dans une moindre mesure sur le Golfe de Guinée. On voit également une influence d’anomalies de TSO positives dans le Pacifique et l’Atlantique sub-tropical, ainsi que la Méditerranée, sur les précipitations sahéliennes.

Lorsque les TSO sont plus chaudes sur le Golfe de Guinée, la convection est accrue sur cette région, et la zone de convergence inter-tropicale (ZCIT) est située plus au sud ; par conséquent, les pluies sahéliennes sont réduites. Cette description est très schématique, car en réalité les interactions entre l’ENSO, les TSO sur le Golfe de Guinée et précipitations dans cette région sont beaucoup plus complexes, et ont évolué au cours des dernières décennies (Janicot et al., 2001; Losada et al., 2012).

Les figures 3.16 (b) et (d) montrent les covariances entre TSO et anomalies de précipitations dans le multi-modèle ENSEMBLES pour le Sahel et la Guinée, respectivement. Le multi-modèle reproduit bien les liens entre les TSO tropicales et les précipitations sur la région du Sahel, ainsi qu’une partie du signal extra-tropical (sur l’Atlantique). Sur la Guinée, le multi-modèle représente correctement l’influence d’anomalies de TSO chaudes dans le Golfe de Guinée sur les précipitations, mais surestime leur importance. Contrairement aux données de référence, le multi-modèle crée un lien entre une anomalie de type La Niña et l’augmentation de précipitations sur le Golfe de Guinée. Ceci peut être lié au décalage trop au sud des précipitations sahéliennes dans la plupart des modèles individuels. En moyenne, les liens entre anomalies de TSO et précipitations sur l’Afrique de l’ouest sont donc représentés de manière assez réaliste dans les modèles, malgré la présence de forts biais de TSO sur les bords d’est dans la plupart des modèles participant au projet ENSEMBLES.

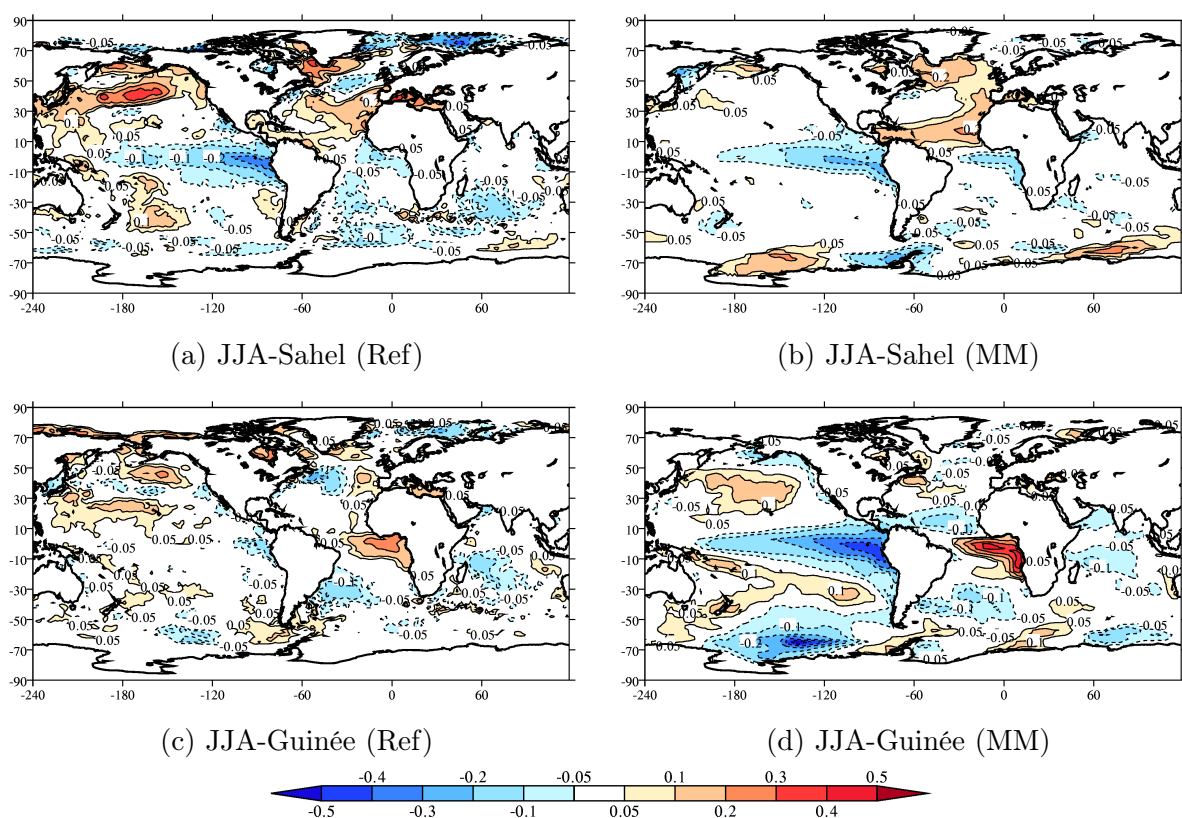


FIGURE 3.16 – Idem figure 3.14, pour les anomalies de précipitations sur le Sahel (en haut) et la Guinée (en bas), dans les données ERA et GPCC (à gauche) et le multi-modèle ENSEMBLES (à droite).

Modèle	mACC		RPSS	
	Guinée	Sahel	Guinée	Sahel
ECMWF	0.173	0.091	-0.004	-0.029
IFMG	0.164	0.179	-0.011	-0.014
MF	0.151	0.118	-0.050	-0.030
UKMO	0.254	0.300	0.006	0.035
CMCC-INGV	0.217	0.023	-0.077	-0.077
MM	0.261	0.240	0.039	0.017

TABLEAU 3.8 – Idem tableau 3.6, pour la saison JAS.

3.2.3 Impact du délai de prévision : performances des prévisions pour la saison JAS

Le tableau 3.8 montre les scores de mACC et RPSS sur les régions de la Guinée et du Sahel pour la saison JAS, correspondant à une prévision de 3 à 5 mois d'échéance. Le contraste entre les scores sur le Sahel et la Guinée est réduit par rapport à JJA, du fait d'une diminution des scores sur le Golfe de Guinée et d'une augmentation des scores sur le Sahel. L'amélioration des résultats sur le Sahel (malgré le décalage de l'échéance d'un mois) peut être due à plusieurs facteurs, comme le décalage dans le temps du déclenchement de la mousson dans les modèles.

La figure 3.17 montre les scores de mACC mois par mois entre juin et septembre. Sur le Sahel, les scores sont meilleurs en août et en septembre qu'au mois de juillet, pour une majorité de modèles ainsi que le multi-modèle. Ceci peut être dû à plusieurs facteurs, comme la mauvaise représentation du saut de mousson dans les modèles, ou encore des erreurs dans la prévision de la variabilité de la date de l'arrivée des précipitations au-dessus de 10°N et donc des cumuls de précipitations en juillet. Sur la Guinée, l'évolution des scores de mACC est quasi-symétrique à celle sur le Sahel, avec un maximum des scores au mois de juillet, puis une dégradation progressive en août et septembre pour la plupart des modèles. Le fait que les scores du mois de juillet soient les plus faibles sur le Sahel et les plus forts sur la Guinée montre qu'il ne s'agit pas uniquement d'un problème de maximum des précipitations restant trop au sud. Ceci a d'ailleurs été confirmé en déplaçant légèrement la frontière que nous avons fixée entre Sahel et Guinée vers le sud, sans grande influence sur les résultats obtenus.

Les covariances d'anomalies en JAS entre TSO et précipitations sur les régions du Sahel et du Golfe de Guinée sont représentées dans la figure 3.18. Les liens dans les références ne changent que très peu par rapport à JJA. La reproduction du lien entre anomalies froides sur le Pacifique tropical et précipitations renforcées sur le Sahel est améliorée dans le multi-modèle, tandis que l'exagération de la covariance entre TSO tropicales et anomalies de précipitations sur le Golfe de Guinée est amplifiée en JAS. Ceci peut en partie expliquer la réduction du contraste des scores entre JJA et JAS, même si la reproduction des covariances n'indique pas nécessairement une bonne performance des modèles en termes de prévision.

Par souci de concision, seules les covariances d'anomalies du multi-modèle sont présentées ici. Nous reviendrons dans le paragraphe suivant sur le cas du modèle couplé du CNRM. La représentation des covariances d'anomalie varie fortement suivant les modèles couplés, certains montrant par exemple une covariance positive entre les TSO du Pacifique Est Tropical et les précipitations sahéliennes. C'est un des intérêts majeurs de l'utilisation du multi-modèle, qui permet d'atténuer les erreurs des modèles individuels et ainsi d'obtenir de meilleures prévisions.

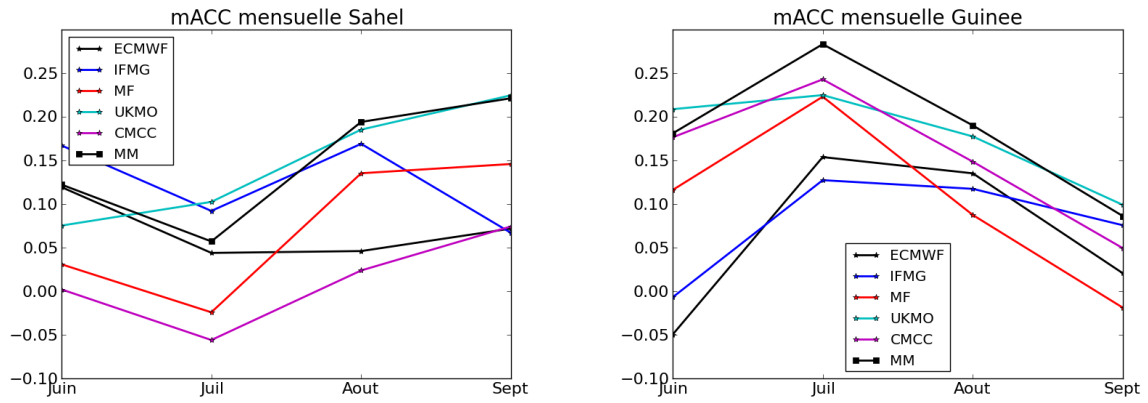


FIGURE 3.17 – Scores mACC calculés pour chaque mois pour les prévisions de précipitations du projet ENSEMBLES sur le Sahel et la Guinée, de juin à septembre 1960–2005.

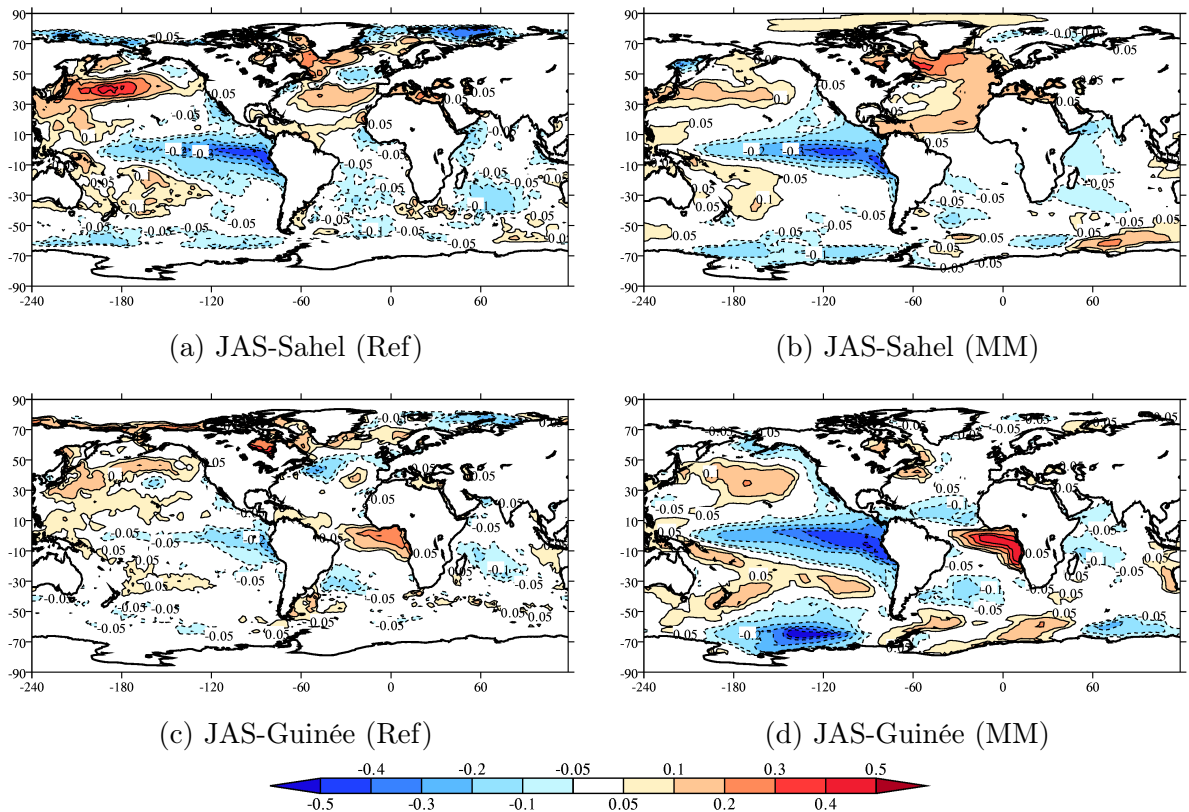


FIGURE 3.18 – Idem figure 3.16, pour la saison JAS.

3.3 Prévisions du modèle CNRM-CM5.1

L'étude des prévisions du projet européen ENSEMBLES a permis d'illustrer l'apport en termes de scores et de représentation des liens observés entre TSO et précipitations de l'approche multi-modèle. Le deuxième jeu de rétro-prévisions saisonnières du projet ENSEMBLES représente l'état de l'art des modèles couplés au milieu des années 2000. Depuis, les différents modèles couplés utilisés dans le projet ont évolué, certaines composantes ont été améliorées, et à la fois les composantes atmosphériques et océaniques ont généralement une plus grande résolution horizontale et/ou verticale. Dans le cas du modèle du CNRM, une nouvelle version a été développée dans le cadre du prochain rapport du GIEC et de l'exercice CMIP5. La partie 2.1 présente les détails du modèle sur lesquels nous ne reviendrons pas ici. Rappelons tout de même l'essentiel des différences entre CNRM-CM5.1 et la version antérieure utilisée pour le projet ENSEMBLES, qu'on notera MF ENS :

- le modèle d'atmosphère ARPEGE passe de la version 4 à la version 5, avec une augmentation de la résolution horizontale d'une troncature TL63 à une TL127 ;
- le modèle d'océan est le modèle NEMO qui remplace le modèle OPA ; la résolution horizontale est de 1° au lieu de 2°, la résolution verticale de 42 niveaux soit 11 niveaux de plus que la version d'OPA utilisée pour ENSEMBLES ;
- la composante « surfaces continentales » est externalisée (modèle SURFEX) dans CNRM-CM5.1 ;
- la glace de mer est prévue par le modèle GELATO, tandis que MF ENS utilisait une glace « if ».²

Ces changements dans le modèle couplé nous incitent à évaluer les évolutions des performances du modèle en prévision saisonnière. Les précipitations africaines sont un étalon intéressant à étudier. Dans ce qui suit, les résultats sont évalués sur la période commune entre les rétro-prévisions ENSEMBLES et les simulations réalisées dans le cadre de cette thèse, à savoir 1979–2005. Les différents ensembles de rétro-prévisions avec CNRM-CM5.1 sont notés CM5 INI, DS RAND et DS SEQ5 selon la méthode de prévision d'ensemble employée (respectivement perturbations initiales, dynamique stochastique aléatoire et séquentielle). On pourra se reporter à la partie 4.3 pour plus de détails. Les précipitations sont cette fois-ci comparées aux réanalyses GPCP, qui existent à l'échelle mensuelle mais aussi quotidienne. Pour les autres champs on utilise toujours les réanalyses ERA-Interim comme référence.

3.3.1 Prévision de la mousson africaine

3.3.1.1 Biais et échelle intra-saisonnière

On s'intéresse ici aux impacts des évolutions du modèle couplé du CNRM sur la représentation de la mousson africaine, en étudiant uniquement la climatologie des précipitations sur l'Afrique de l'ouest et non pas la prévision de la variabilité interannuelle.

La figure 3.19 montre les biais des précipitations pour JJA 1979–2005 par rapport à la base de données GPCP sur la région, pour les ensembles MF ENS et DS SEQ5. Les flèches quantifient les directions et intensités des biais de vent par rapport à la base de données ERA-Interim, à 10 mètres d'altitude pour MF ENS, et 925 hPa pour la simulation avec CNRM-

2. Paramétrisation consistant à introduire une couche de glace fictive si la température de l'océan atteint la température de congélation, afin de reproduire les valeurs flux océan-atmosphère typiquement observées.

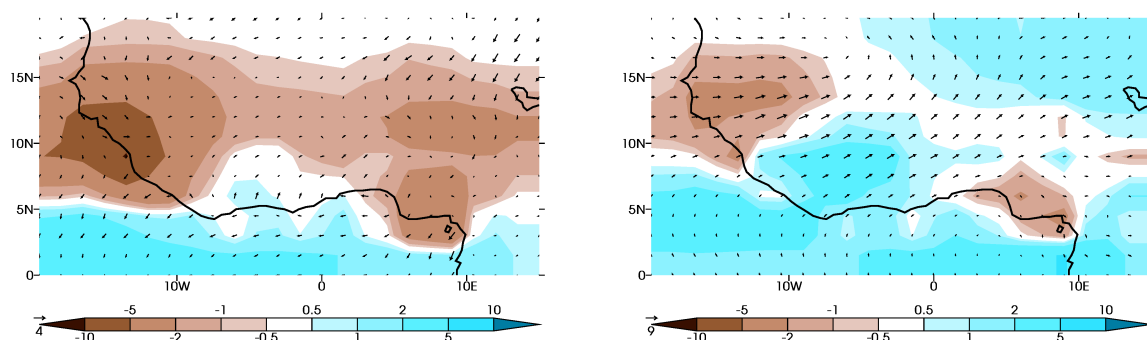


FIGURE 3.19 – Biais des précipitations (couleurs) et de vent près de la surface en JJA pour les rétro-prévisions du modèle du CNRM dans ENSEMBLES (MF ENS, à gauche) et de CNRM-CM5.1 avec dynamique stochastique séquentielle (DS SEQ5, à droite ; voir la partie 4.3 du chapitre suivant). Les champs de vent sont le vent à 10 mètres pour MF ENS et le vent à 925 hPa pour DS SEQ5.

CM5.1. L'ensemble MF ENS est trop sec sur le plupart des points de terre, avec un déficit plus marqué encore sur l'ouest du domaine. Au-dessus du Golfe de Guinée les pluies sont excédentaires. Un biais de vent amène de l'air sec saharien sur l'est du domaine, induisant des déficits de précipitations de 1 à 5 mm par jour. Pour l'expérience DS SEQ5 avec CNRM-CM5.1, les biais de vent sont orientés différemment, et amplifient les vents climatologiques sur la région. Ces vents amènent au contraire de l'air humide du Golfe de Guinée vers l'intérieur des terres, ce qui engendre un excédent de précipitations sur la plupart du domaine. On retrouve deux zones de déficit de précipitations présentes avec MF ENS : le Sénégal ainsi que le sud du Nigeria.

Les données GPCP permettent d'étudier ces différences à l'échelle intra-saisonnière. On représente l'évolution temporelle climatologique de la mousson de mai à octobre à l'aide de moyennes longitudinales entre 10°W et 10°E des pentades glissantes de précipitations. Ces diagrammes de Hovmöller sont montrés à la figure 3.20, pour les données quotidiennes GPCP, MF ENS et l'ensemble CNRM-CM5.1 avec perturbations initiales. Les défauts visibles lors de l'étude des biais saisonniers sont en partie expliqués par cette figure : pour MF ENS, la bande de précipitations plus marquées atteint à peine 10°N, et les cumuls quotidiens de précipitations sont trop faibles sur la plupart des latitudes tout au long de la saison de mousson, sauf au mois de mai près de l'équateur. Avec CNRM-CM5.1, les précipitations sont plus marquées et même surestimées sur une grande partie du domaine, ce qui est cohérent avec les biais de précipitations de la figure 3.19. Par contre, la position du maximum de précipitations au cours de la saison ainsi que le déclin de la mousson en septembre et octobre paraît plus réaliste qu'avec MF ENS. Les mêmes diagrammes ont permis de comparer les versions avec et sans dynamique stochastique des prévisions avec CNRM-CM5.1 sur la saison MJJA, sans montrer de différence majeure.

3.3.1.2 Téléconnexions avec les TSO

La figure 3.21 (page 98) montre les covariances d'anomalies entre TSO et précipitations sur le Sahel et la Guinée pour les données des réanalyses GPCP et ERA-Interim, et les différents jeux d'expériences étudiés dans cette partie.

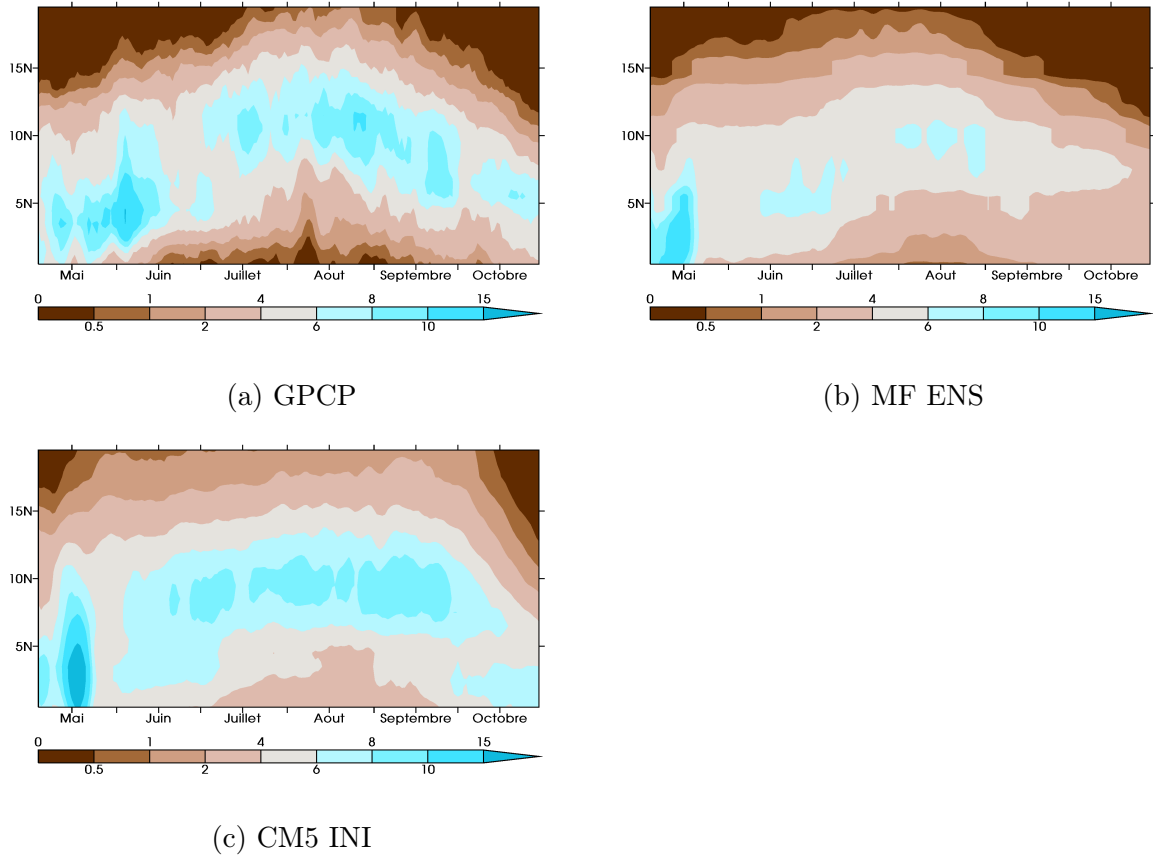


FIGURE 3.20 – Diagrammes Hovmöller des précipitations quotidiennes (moyennées sur des pentades glissantes) de MJJASO pour (de haut en bas) GPCP, MF ENS et CM5 INI. Les moyennes sont sur 1997–2005 pour MF ENS, 1997–2010 pour GPCP et CM5 INI.

En comparant les figures 3.16 (a) et (c) aux figures 3.21 (a) et (b), on note certaines différences entre les liens de TSO et anomalies de précipitations sur ces régions entre les périodes 1960–2005 et 1979–2005. Sur la Guinée, l’influence des TSO est plus forte au cours de la deuxième période, avec l’apparition d’un lien entre une anomalie de type La Niña et des précipitations accrues sur le Golfe de Guinée. En ce qui concerne le Sahel, la différence la plus évidente est le changement de signe des anomalies de TSO sur le Golfe de Guinée liées à des précipitations accrues sur la région. Ces résultats sont cohérents avec ceux montrés par les travaux de Losada et al. (2012). Les auteurs calculent l’évolution au cours du vingtième siècle des corrélations entre les TSO sur les régions Niño 3 et Atl 3 et les précipitations sur le Golfe de Guinée et le Sahel pour des fenêtres glissantes de 20 ans. Ils montrent que la corrélation négative entre les TSO Atl 3 et le Sahel n’est plus significative à partir des années 1970, et commence à changer de signe à la fin du siècle. Ils constatent également l’apparition d’une corrélation négative significative entre les TSO Niño 3 et les précipitations sur le Golfe de Guinée à partir de la même période. Le seul lien stationnaire entre TSO et précipitations sur l’Afrique de l’ouest est la corrélation positive significative entre TSO sur la boîte Atl 3 et précipitations sur le Golfe de Guinée.

Comment les différentes versions du modèle couplé du CNRM représentent-elles ces liens sur la période 1979–2005 ? La différence la plus frappante entre la version ultérieure du modèle

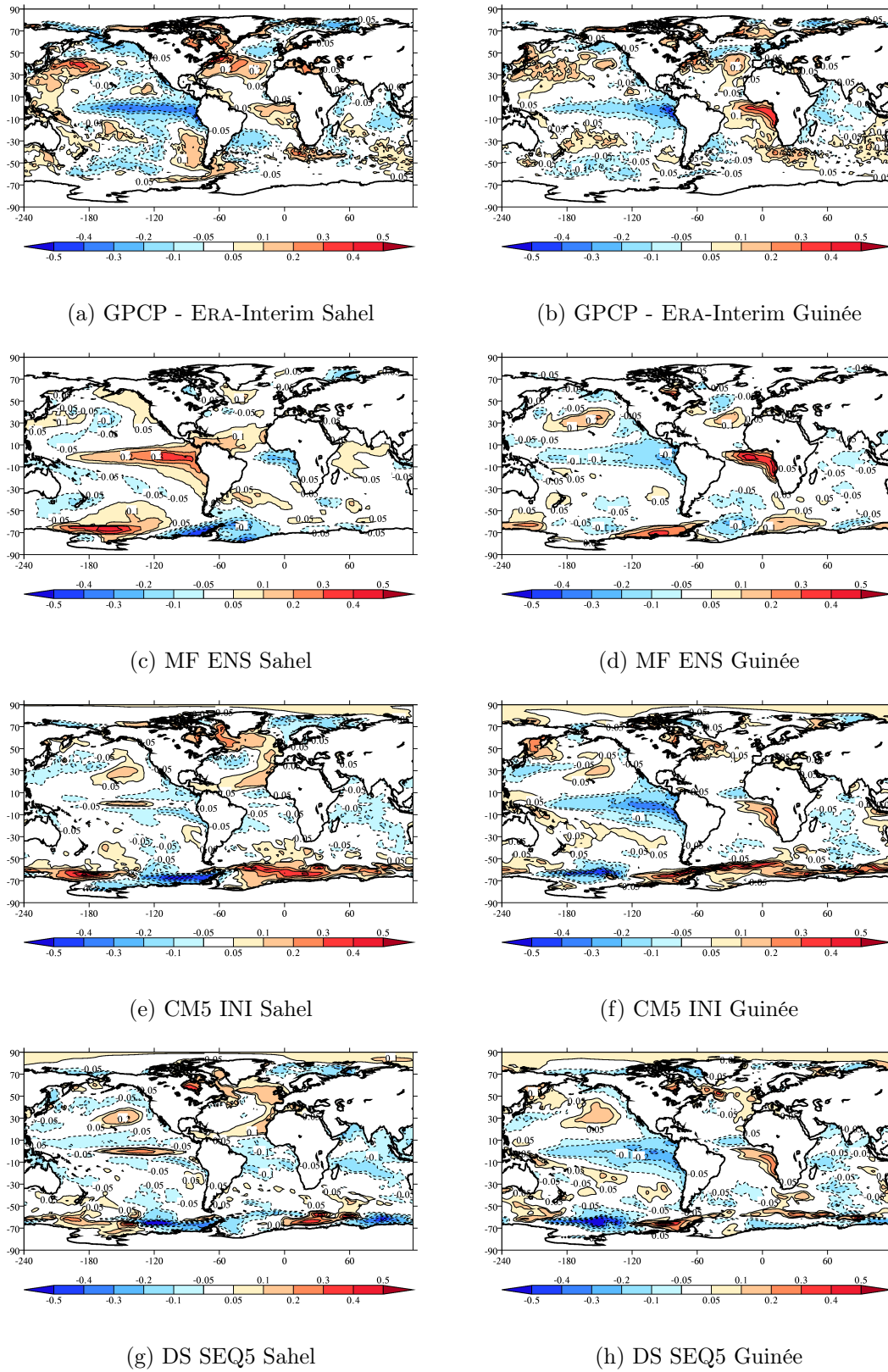


FIGURE 3.21 – Idem figure 3.16, mais sur la période 1979–2005 pour les données GPCP (précipitations) et ERA-Interim (TSO), et les ensembles MF ENS, CM5 INI et DS SEQ5.

Modèle	Afrique de l'ouest			Sahel			Guinée		
	5%	95%	ACC	5%	95%	ACC	5%	95%	ACC
MF ENS	-	-	0.21	-	-	0.04	-	-	0.30
CM5 INI	0.14	0.23	0.19	-0.03	0.12	0.05	0.20	0.32	0.26
DS RAND	0.13	0.22	0.18	-0.01	0.11	0.06	0.18	0.27	0.23
DS SEQ5	0.16	0.28	0.23	0.00	0.10	0.05	0.22	0.38	0.30

TABEAU 3.9 – ACC moyenne (moyenne de la transformée en Z des séries d'ACC spatiales pour JJA 1979–2005) sur les régions de l'Afrique de l'ouest pour le modèle de Météo-France du projet ENSEMBLES et plusieurs ensembles utilisant CNRM-CM5.1. Pour les ensembles de CNRM-CM5.1 (15 membres), on montre une fourchette de 5% à 95% et la moyenne des ACC avec 9 membres tirés aléatoirement parmi les 15.

et CNRM-CM5.1 est la représentation des liens entre TSO sur le Pacifique et anomalies de précipitations sur le Sahel. En effet, sur la figure 3.21 (c), l'ensemble MF ENS inverse le signe des anomalies de TSO, en liant les précipitations sahéliennes accrues à une structure de type El Niño ainsi qu'à des anomalies froides de TSO sur le Golfe de Guinée. Ces défauts sont en grande partie atténués dans les ensembles CNRM-CM5.1, que ce soit avec ou sans introduction de la dynamique stochastique. Pour les précipitations sur le Golfe de Guinée, il y a peu de changements entre les différentes versions du modèle, si ce n'est dans l'intensité des covariances d'anomalies de TSO. Globalement, la représentation des liens entre TSO et précipitations ouest-africaines est légèrement améliorée dans la version CNRM-CM5.1 du modèle couplé du CNRM.

3.3.1.3 Scores des prévisions de JJA

Comment les évolutions de la représentation de la mousson africaine se traduisent-elles en termes de scores? Pour répondre à cette question, les corrélations d'anomalies de précipitations ainsi que le RPSS sont calculés sur la période commune des rétro-prévisions. Pour les corrélations d'anomalies, le score calculé est le zACC présenté dans la partie 2.2.2.

Les ensembles de rétro-prévisions saisonnières avec CNRM-CM5.1 ayant été réalisés avec 15 membres au lieu de 9, on peut calculer des fourchettes de scores en tirant aléatoirement 9 membres parmi 15 un grand nombre de fois. Cette méthode permet d'estimer l'incertitude liée à la taille des ensembles et de donner une indication quant à la significativité des différences de scores entre deux ensembles. Les résultats sont montrés dans le tableau 3.9. Malgré les changements au niveau de la représentation des liens avec les TSO, les scores de corrélation des anomalies demeurent très proches de ceux obtenus avec la version utilisée dans le cadre du projet ENSEMBLES. Le score de MF ENS est toujours compris dans la fourchette des scores des versions CNRM-CM5.1, à l'exception du score sur la Guinée pour l'ensemble DS RAND qui est légèrement moins bon.

De manière générale, les scores obtenus avec des ensembles de 15 membres sont meilleurs que ceux obtenus avec 9 membres. On peut se demander dans quelle mesure les scores d'ACC dépendent de la taille des ensembles, et si l'amélioration des scores est possible en utilisant des ensembles plus grands. Pour étudier ceci, trois ensembles de 15 membres de prévisions similaires (CNRM-CM5.1 avec dynamique stochastique) sont regroupés, et un grand nombre de tirages aléatoires de n membres parmi les 45 est effectué. La moyenne des ACC obtenus

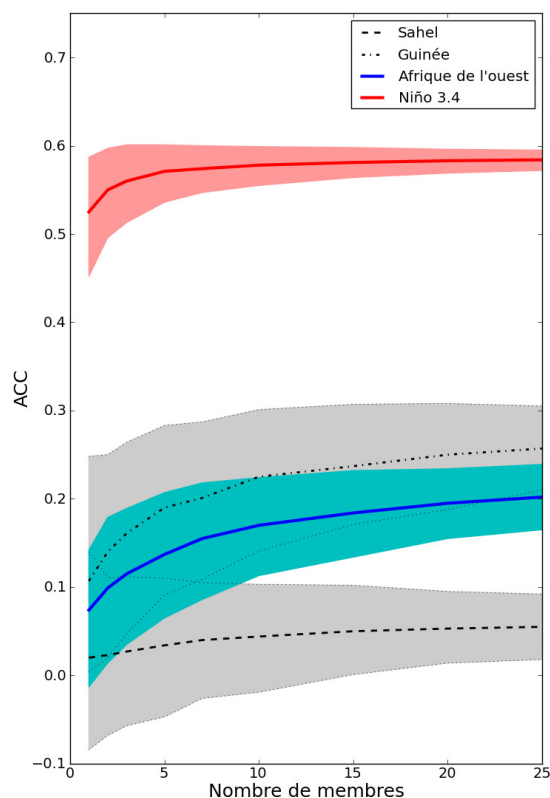


FIGURE 3.22 – Influence de la taille des ensembles sur l'ACC des prévisions de précipitations. Les membres des expériences DS RAND, DS SEQ5 et DS N34 sont regroupés en un grand ensemble de 45 membres, et n membres sont tirés de manière aléatoire. Les résultats montrés sont pour n de 1 à 25 membres parmi 45, pour un grand nombre de tirages. Ceci permet d'estimer la moyenne des transformées en Z des ACC spatiales (traits gras) ainsi que la fourchette de 5%-95% de ce score pour les régions de l'Afrique de l'ouest et la région Niño 3.4.

ainsi que la fourchette des scores de 5% à 95% sont montrés pour les régions de l'Afrique de l'ouest (figure 3.22). A titre de comparaison, la figure présente également les résultats obtenus sur la région Niño 3.4.

Ces résultats permettent de constater que les fourchettes de scores de précipitations sur l'Afrique de l'ouest sont nettement plus larges que celles sur la région Niño 3.4, et que si l'ACC sature au-delà d'une dizaine de membres dans le cas de la boîte Niño 3.4, l'augmentation de la taille des ensembles permet encore d'améliorer l'ACC sur les régions de l'Afrique de l'ouest et de la Guinée, et dans une moindre mesure sur le Sahel. D'ailleurs, sur cette région, le 95^{ème} centile des scores décroît avec la taille des ensembles, même si dans tous les cas ces scores restent très faibles. Pour les trois régions ces courbes semblent s'approcher d'une asymptote horizontale, ce qui signifie que sans amélioration du modèle on ne peut espérer accroître les scores simplement en démultipliant la taille de l'ensemble.

Les scores de RPSS pour les terciles de précipitations sont donnés dans le tableau 3.10. Ils sont très marginalement meilleurs avec CNRM-CM5.1 que la version du projet ENSEMBLES, pour laquelle la prévision probabiliste est moins bonne que l'utilisation de la climatologie. On remarque que la fiabilité est améliorée tandis que le terme de résolution, qu'on veut le plus élevé possible, est légèrement diminué. Il y a peu d'impact de la dynamique stochastique sur les scores obtenus. Dans l'article publié dans *Tellus A*, les valeurs de RPSS étaient plus importantes sur le Golfe de Guinée, même si on trouvait des RPSS positifs sur certaines zones du Sahel (voir figure 3.7). C'est le cas avec le modèle MF ENS, et également avec les différents ensembles utilisant CNRM-CM5.1.

Ensemble	Rel/Unc	Res/Unc	RPS	RPSS
Climatologie	0.	0.	0.222	0.
MF ENS	0.491	0.465	0.228	-0.026
CM5 INI	0.439	0.443	0.221	0.004
DS RAND	0.437	0.441	0.221	0.005
DS SEQ5	0.427	0.432	0.221	0.005

TABLEAU 3.10 – Scores de RPSS et décomposition en fiabilité, Rel, et résolution, Res (divisés par le terme d’incertitude Unc), pour les prévisions des précipitations en JJA sur l’Afrique de l’ouest des différentes versions du modèle couplé du CNRM.

Modèle	DJF Af. du sud			MAM GHA			SON GHA		
	5%	95%	ACC	5%	95%	ACC	5%	95%	ACC
MF ENS	-	-	-0.02	-	-	0.	-	-	0.36
CM5 INI	-0.02	0.07	0.03	-0.04	0.04	0.	0.20	0.28	0.24
DS RAND	0.04	0.16	0.10	-	-	-	-	-	-
DS SEQ5	0.10	0.21	0.15	0.03	0.13	0.08	0.24	0.34	0.29

TABLEAU 3.11 – Idem tableau 3.9, mais pour les régions d’Afrique du sud en été austral (DJF) et la Corne de l’Afrique pendant les saisons des « long rains » (MAM) et des « short rains » (SON).

Malgré certaines améliorations constatées dans la représentation des liens entre TSO et précipitations et la localisation et l’intensité des précipitations, les scores des prévisions saisonnières sont assez similaires entre les deux versions du modèle couplé du CNRM étudiées ici.

3.3.2 Le cas des autres régions d’Afrique

L’étude précédente est étendue dans cette sous-partie aux cas des autres régions d’Afrique étudiées dans l’article de *Tellus A*.

3.3.2.1 Prévision des précipitations sud-africaines

Si les scores du multi-modèle du projet ENSEMBLES étaient assez bons sur l’Afrique du sud pour les précipitations de l’été austral (DJF), ceux du modèle MF ENS étaient les moins bons des cinq modèles individuels. MF ENS ne montrait en moyenne aucune corrélation des anomalies de précipitations, et on retrouve ces résultats dans le tableau 3.11 pour la période 1979–2005. Dans ce cas, l’apport du modèle CNRM-CM5.1 et de l’introduction de la dynamique stochastique est évident. Les fourchettes de scores obtenues avec CNRM-CM5.1 se détachent du score de MF ENS, et la dynamique stochastique séquentielle permet de considérablement améliorer les scores d’ACC par rapport aux perturbations initiales.

En termes de RPSS, le modèle CNRM-CM5.1 avec dynamique stochastique améliore à la fois les composantes de fiabilité et de résolution, ce qui se traduit par une nette amélioration du score de RPSS moyen sur le domaine (tableau 3.12). Outre le score moyen, c’est aussi le nombre de points de grille du domaine où le RPSS est positif qui augmente avec les nouvelles expériences. Pour MF ENS, seuls 32% des points de grille ont un RPSS positif, tandis que

Ensemble	Rel/Unc	Res/Unc	RPS	RPSS
MF ENS	0.468	0.430	0.231	-0.038
CM5 INI	0.453	0.438	0.226	-0.015
DS RAND	0.448	0.450	0.222	0.002
DS SEQ5	0.425	0.437	0.220	0.012

TABEAU 3.12 – Idem tableau 3.10, pour l'Afrique du sud en DJF.

	GHA MAM				GHA SON			
Ensemble	Rel/U	Res/U	RPS	RPSS	Rel/U	Res/U	RPS	RPSS
MF ENS	0.433	0.416	0.226	-0.017	0.387	0.397	0.220	0.010
CM5 INI	0.448	0.428	0.227	-0.020	0.433	0.436	0.222	0.003
DS SEQ5	0.456	0.443	0.225	-0.013	0.446	0.464	0.218	0.018

TABEAU 3.13 – Idem tableau 3.10, pour la Corne de l'Afrique en MAM (à gauche) et SON (à droite).

pour les trois ensembles de prévisions avec CNRM-CM5.1, cette proportion passe à 37%, 50% et 66% respectivement pour INI, DS RAND et DS SEQ5.

3.3.2.2 « Short rains » et « long rains » de la Corne de l'Afrique

Sur la région de la Corne de l'Afrique, les résultats sont plus mitigés. La saison des « long rains » (MAM), pour laquelle les modèles du projet ENSEMBLES ne montraient aucun résultat probant, pose toujours problème dans la nouvelle version du modèle. Le tableau 3.11 illustre ce point : les scores de zACC restent insignifiants et même si ceux obtenus avec l'ensemble DS SEQ5 paraissent meilleurs, la fourchette de zACC n'est pas disjointe de celle de CM5 INI.³ En termes de RPSS, l'apport est également nul, comme le montrent les scores du tableau 3.13. Cette saison des pluies reste pour l'heure très problématique en prévision saisonnière avec des modèles dynamiques couplés. Une étude récente de Camberlin et al. (2009) met en lumière certaines sources potentielles d'erreurs des modèles : les « long rains » ont moins de cohérence spatiale et temporelle que les « short rains » dans la date d'arrivée des pluies, l'intensité des précipitations, le nombre de jours de pluie et la fin de la saison des pluies. Un accroissement de la dispersion d'ensemble ne suffit pas dans ce cas à améliorer les prévisions.

Pour la saison des « short rains » (SON), les scores d'ACC sont moins bons avec les ensembles CNRM-CM5.1 que pour l'ensemble MF ENS, dont l'ACC sur la région dépasse le 95^{ième} centile de DS SEQ5, qui est lui-même meilleur que CM5 INI. L'apport de la dynamique stochastique est visible pour les scores de RPSS, la résolution augmente plus que le terme de fiabilité par rapport à MF ENS, ce qui donne un RPSS plus élevé.

3. Seul l'ensemble avec dynamique stochastique DS SEQ5 a été implémenté pour les saisons intermédiaires MAM et SON.

3.4 Synthèse

L'objectif de ce chapitre était d'évaluer l'état de l'art des modèles dynamiques couplés en prévision saisonnière sur plusieurs régions d'Afrique, où les attentes en termes d'apport de la prévision saisonnière de précipitations sont multiples. La prévision saisonnière sur ces régions sert d'aide à la prise de décision pour des questions de santé publique, d'énergie et d'agriculture. Dès lors, il est intéressant et important de caractériser et quantifier les réussites et faillites des modèles de climat sur ces régions.

L'article publié dans *Tellus A* a permis de mettre en évidence les avantages d'une approche multi-modèle. Les scores du multi-modèle du projet européen ENSEMBLES montrent qu'il est possible d'avoir des prévisions saisonnières de précipitations apportant une information plus utile sur les différentes régions que l'utilisation de la climatologie. L'utilisation de prévisions d'ensemble probabilistes permet de donner une information plus complète sur les caractéristiques d'une saison humide. L'approche multi-modèle permet surtout de réduire les erreurs de modèles individuels, sachant qu'aucun modèle du projet n'est systématiquement meilleur que les autres sur les régions d'Afrique choisies dans l'étude. Une autre conclusion majeure de l'article est l'importance de la bonne représentation des liens entre TSO et anomalies de précipitations, condition nécessaire, mais non suffisante, pour avoir de bons scores de prévision. Les résultats obtenus sur les différentes régions confirment par ailleurs que les conditions océaniques tropicales sont une source déterminante de prévisibilité pour les modèles couplés, puisque les meilleurs résultats sont obtenus au-dessus des régions où ces téléconnexions sont les plus marquées.

L'étude plus détaillée sur l'Afrique de l'ouest confirme ces derniers points, puisqu'elle illustre les écarts de performances des prévisions saisonnières sur la région du Golfe de Guinée avec le Sahel. Malgré l'existence de forts biais aux bords est-océaniques dans la plupart des modèles, le multi-modèle reproduit bien les covariances d'anomalies entre TSO et précipitations sur le Sahel. Les scores sur cette région n'en demeurent pas moins faibles, et on montre que ceux-ci sont nettement meilleurs les années de TSO plus élevées sur la boîte Niño 3.4 que pour les années de TSO normales ou fraîches. Ceci souligne également la non-linéarité de la réponse des précipitations sahéliennes aux TSO tropicales.

Dans ce chapitre, nous nous sommes également attachés à évaluer les impacts des évolutions du modèle couplé du CNRM sur la qualité des prévisions saisonnières de précipitations sur l'Afrique. Le modèle couplé CNRM-CM5.1 ne présente plus les mêmes biais que son prédécesseur utilisé dans le projet ENSEMBLES, et l'évolution à l'échelle intra-saisonnière des précipitations de la mousson ouest-africaine est plus réaliste et mieux localisée. De plus, les covariances d'anomalies de TSO et précipitations sur le Sahel, qui étaient inversées par rapport aux données des réanalyses dans la version du projet ENSEMBLES, sont en partie corrigées. Ces changements ne se traduisent pas par des progrès en termes de scores, si ce n'est pour les scores probabilistes. Toutefois, ces conclusions ne sont pas valables sur toutes les régions d'Afrique étudiées. Sur l'Afrique du sud, la version 5 du modèle couplé est significativement meilleure que celle utilisée dans ENSEMBLES. Dans ce cas, l'introduction d'une dynamique stochastique améliore encore les résultats, comme c'est le cas sur d'autres régions tempérées. Les résultats sur la Corne de l'Afrique sont plus décevants, dans la mesure où aucune amélioration notable des scores n'est observée avec CNRM-CM5.1 pour la saison des « long rains », tandis que les scores sont équivalents ou légèrement moins bons que dans le projet ENSEMBLES pour la saison des « short rains ».

Il pourrait être tentant de voir dans ses résultats la preuve que la prévision des saisons des pluies sur ces régions d'Afrique à l'échelle saisonnière est illusoire. Cependant, les évolutions des modèles couplés, et les améliorations des scores des modèles individuels en utilisant une approche de type multi-modèle, montrent au contraire qu'il faut persévérer dans cette direction. L'avantage des modèles dynamiques par rapport à des approches de type statistique est leur indépendance vis-à-vis des évolutions du climat. A titre d'exemple, les modifications des liens entre TSO et précipitations sahéliennes observées au cours de la deuxième moitié du vingtième siècle seraient difficiles voire impossibles à prendre en compte dans un modèle statistique, et réduiraient sa robustesse. Le défi consiste à améliorer la représentation des composantes du climat dans les modèles couplés afin de tirer parti des sources de prévisibilité là où elles existent. Des approches de type « model output statistics » (MOS) permettent de tirer parti des points forts des prévisions des modèles couplés. Ainsi, Philippon et al. (2010) et Ndiaye et al. (2011) ont montré que les modèles dynamiques pouvaient donner lieu à des prévisions saisonnières de la mousson africaine améliorées, en exploitant les TSO et la circulation près de la surface ou à 200 hPa de différents modèles couplés dans le cadre de prévisions statistico-dynamiques.

Chapter Summary

The main aim of this chapter is to assess the state of the art in dynamical coupled model seasonal predictions of precipitation over several regions of Africa. Seasonal forecasts over these regions are useful for decision-making in areas such as public health, energy or agriculture. It is therefore important and relevant to quantify the successes and flaws of climate models over these regions.

The article published in Tellus A, presented in the first part of this chapter, shows how useful a multi-model approach can be. The ENSEMBLES project multi-model scores prove that seasonal precipitation forecasts can give more useful information than climatology over the regions of interest. Using probabilistic ensemble forecasts helps give a more thorough outlook on rainy season characteristics. The multi-model method is efficient in reducing individual model errors, and no single model is systematically better than the others over each region studied. Another main conclusion in the article is that representing links between SST and precipitation anomalies is a necessary step towards improving forecast skill. Strong anomalies in tropical ocean conditions are a possible source of predictability in coupled models, since better results are found over regions where precipitation teleconnections with SST anomalies are the strongest.

A more detailed study over West Africa is consistent with these conclusions, and exhibits gaps in seasonal forecast skill between the Gulf of Guinea region and the Sahel. Despite strong biases over the eastern coasts of the oceans in most models, the multi-model represents anomaly covariances between SST and sahelian precipitation quite well. Scores over the region remain low, and are clearly better when SSTs are higher over the Niño 3.4 region than when SSTs are normal or below-normal. This illustrates the fact that the Sahel precipitation response to tropical SST anomalies is non-linear.

In this chapter, we also focused on the evaluation of improvements and setbacks between the former version of the CNRM climate model used in the ENSEMBLES project and the new coupled model CNRM-CM5.1. The latter exhibits a more realistic intra-seasonal evolution of zonal precipitation maxima during the West African monsoon season. Anomaly covariance between SST and precipitation over the Sahel is also somewhat corrected. Yet, these changes do not lead to improvements in terms of scores, save for some probabilistic skill scores. These conclusions do not hold over other regions of Africa. Over South Africa in DJF, CNRM-CM5.1 is significantly better than the version used in ENSEMBLES. In this case, the introduction of stochastic dynamics further improves the results, as over other mid-latitude regions. Results over the Horn of Africa are disappointing, in so far as no improvement in scores is found with CNRM-CM5.1 for the “long rains” season whereas scores are equal to or slightly lower than the ENSEMBLES version ensemble for the “short rains” season.

It could be tempting to see in these results undeniable proof that seasonal predictions of precipitation over these regions of Africa is impossible, yet further work could lead to improvements. Combinations of dynamical modeling and statistical methods, such as model output statistics (MOS), can help extract useful information from dynamical seasonal predictions over these regions, as shown by Philippon et al. (2010) and Ndiaye et al. (2011).

TROISIÈME PARTIE

LA DYNAMIQUE STOCHASTIQUE

PERTURBATIONS STOCHASTIQUES DE LA DYNAMIQUE DU MODÈLE ARPEGE-CLIMAT

Le chapitre précédent a permis d'illustrer l'état de l'art actuel des modèles de circulation générale couplés en prévision saisonnière à l'aide d'un exemple précis (et particulièrement délicat à prévoir), celui des précipitations sur l'Afrique. Un aspect fondamental en prévision saisonnière est l'amélioration des prévisions des modèles par la prise en compte des incertitudes liées à la modélisation. L'approche multi-modèle est en cela une référence, et l'exemple le plus probant de ces améliorations.

Dans ce chapitre, nous allons explorer une autre approche pour prendre en compte les incertitudes liées à la formulation d'un seul modèle couplé. La méthode présentée ici appartient à une catégorie de méthodes ensemblistes couramment appelée « physique stochastique », mais dans le cadre de cette thèse, elle est baptisée « dynamique stochastique » au vu des choix d'implémentation qui sont faits. Si de nombreux centres de recherche ont d'ores et déjà exploré plusieurs techniques de perturbations stochastiques des modèles d'atmosphère pour différentes échelles temporelles et spatiales de prévision, la méthode développée ici est (à notre connaissance) unique, dans la mesure où elle intègre à la fois une technique s'apparentant à la correction de flux, et un aspect aléatoire dans les perturbations. L'objet de ce chapitre est de présenter en détail cette méthode, de montrer les principaux résultats obtenus en prévision saisonnière, et d'affiner la définition de cette dynamique stochastique en étudiant les impacts de certains choix de perturbations (fréquence, échelle de temps).

La première partie de ce chapitre présente les détails de la méthodologie définissant la dynamique stochastique, et l'implémentation avec CNRM-CM5.1. Dans un premier temps, les perturbations ajoutées à la dynamique du modèle ARPEGE-Climat ont été définies à l'aide d'une technique appelée nudging (ou guidage) d'anomalie. Le développement de cette première méthode, et les résultats sur la période de réanalyse 1989–2010 pour les saisons DJF et JJA font l'objet de la deuxième partie. Ensuite, une alternative au nudging d'anomalie, consistant à dériver les perturbations du modèle avec un nudging itéré deux fois, a permis d'obtenir des résultats plus probants pour la saison DJF. Cette deuxième méthode est présentée en

détail dans la troisième partie de ce chapitre, où les rétro-prévisions sont évaluées sur une période plus longue (DJF et JJA 1979–2010). Dans cette dernière partie, les deux méthodes sont comparées sur la période commune. Plusieurs variantes à une méthode de perturbations aléatoires à haute fréquence sont ensuite étudiées, afin de mieux comprendre les différentes sources des impacts de la dynamique stochastique sur les rétro-prévisions.

4.1 Théorie et méthodologie retenue

Cette partie aborde les différentes étapes du raisonnement scientifique à l’origine du développement de la méthode de dynamique stochastique mise en place dans le cadre de cette thèse. La technique s’appuie sur plusieurs travaux décrits au chapitre 1, portant sur les méthodes stochastiques et la correction des erreurs des modèles de climat.

4.1.1 De la théorie à l’implémentation dans le modèle

La méthode retenue dans le cadre de cette thèse est d’introduire des perturbations stochastiques additives dans les équations de la dynamique du modèle ARPEGE-Climat. Le choix de perturbations additives plutôt que multiplicatives est très courant dans le cadre des modèles de circulation générale. En ce qui concerne le modèle ARPEGE, ce choix permet d’implémenter une dynamique stochastique reposant sur une structure de code informatique préexistante, servant pour le guidage du modèle (nudging). Au contraire, introduire des perturbations multiplicatives nécessiterait de reformuler l’ensemble des équations différentielles modélisées sous une forme stochastique, ce qui représente un travail colossal. En modélisation du climat, les perturbations stochastiques multiplicatives sont donc rencontrées essentiellement dans des schémas de convection ou d’autres paramétrisations physiques, ou alors étudiées avec des modèles théoriques simplifiés comme le modèle Lorenz 96.

L’idée ici est d’allier des perturbations de la dynamique du modèle aux corrections de certaines erreurs de celui-ci, c’est à dire de renforcer le lien entre les travaux portant sur les perturbations stochastiques dans les modèles de circulation générale (voir par exemple Buizza et al., 1999; Weisheimer et al., 2011) et les techniques de correction d’erreurs systématiques des modèles explorées dans le cadre du projet POTENTIALS (voir D’Andrea et Vautard, 2000; Guldberg et al., 2005). Pour ce faire, on va introduire des perturbations tirées aléatoirement parmi un ensemble de corrections d’erreurs de tendance de la composante ARPEGE-Climat de CNRM-CM5.1 en mode prévision saisonnière.

Le choix de la méthode de perturbation du modèle n’est pas anodin, et est en partie contraint par la formulation de celui-ci. Le modèle ARPEGE-Climat utilisé dans cette thèse inclut non seulement la possibilité d’activer le nudging de variables pronostiques du modèle, mais également celle d’ajouter aux variables pronostiques des termes bruts contenus dans un fichier de données. Cette ossature de code a servi à l’implémentation des perturbations stochastiques additives. Les variables pronostiques du modèle sont perturbées en ajoutant des champs spectraux de perturbation. Les possibilités en matière de variables à perturber sont donc restreintes à l’humidité spécifique, la température, la vorticité et la divergence des champs de vents, à tous les niveaux verticaux d’ARPEGE-Climat.

Dans le cadre de cette thèse, seule la partie rotationnelle du vent est perturbée, afin

d'éviter de générer localement des perturbations de l'équilibre géostrophique, et de favoriser les perturbations à plus grande échelle. Les variables perturbées sont donc la température, la vorticité et l'humidité.

La dynamique stochastique ainsi définie peut s'écrire par l'équation 4.1 où \mathbf{X} désigne les variables pronostiques du modèle qu'on perturbe, $\mathbf{M}(\mathbf{X}(t), t)$ l'incrément de ces variables par le modèle non perturbé, et $\delta\mathbf{X}_t$ le terme stochastique à l'instant t . Cette équation est commune à toutes les méthodes de perturbations stochastiques additives. Les spécificités de la méthode développée dans le cadre de cette thèse sont à chercher dans la définition du terme $\delta\mathbf{X}_t$.

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \mathbf{M}(\mathbf{X}(t), t) + \delta\mathbf{X}_t \quad (4.1)$$

Dans cette méthode de dynamique stochastique, $\delta\mathbf{X}_t$ est tiré de manière aléatoire parmi un ensemble $\{\delta\mathbf{X}\}$ de champs spectraux globaux de perturbations. On change de $\delta\mathbf{X}$ à chaque période \tilde{t} . Les termes $\delta\mathbf{X}$ correspondent à des erreurs de tendance initiale du modèle estimées *a priori*. Telle que la méthode est définie ici, afin de tirer aléatoirement une perturbation $\delta\mathbf{X}$, on tire une date d'une période d'intégration du modèle pendant laquelle les champs de perturbations ont été créés. Le $\delta\mathbf{X}$ ainsi déterminé correspond à une correction d'erreur de tendance cohérente pour les différents champs T , Q et Ψ , et tous les niveaux verticaux du modèle.

Les perturbations sont faites sur tous les niveaux verticaux sauf à la surface. Un profil vertical d'atténuation des perturbations sur les niveaux les plus bas du modèle ARPEGE-Climat est introduit afin d'éviter toute incohérence aux interfaces avec les autres composantes de CNRM-CM5.1.

4.1.2 Erreurs de tendance initiale du modèle ARPEGE-Climat

La première étape de la mise en œuvre de la dynamique stochastique est de générer un ensemble de perturbations $\{\delta\mathbf{X}\}$. Ces termes doivent correspondre à des corrections d'erreur de tendance initiale que le modèle fait en cours de prévision. Afin d'estimer ces erreurs de tendance initiale, la méthode employée par Kaas et al. (1999) et Guldberg et al. (2005) est retenue. Elle repose sur la technique de nudging (ou guidage) du modèle vers des données de référence (Jeuken et al., 1996). Le paragraphe 2.3.1 donne plus de détails sur le nudging. Les $\{\delta\mathbf{X}\}$ comprennent des termes de rappel du modèle vers un guide, obtenus lors d'une première intégration du modèle sur une période donnée. Les ensembles $\{\delta\mathbf{X}\}$ peuvent ensuite être échantillonnés, de manière à tirer les perturbations dans des populations correspondant à un mois ou une saison donnée, en prenant soin de ne pas tirer un terme de correction correspondant à la saison courante dans le cas des rétro-prévisions.

Rappelons l'équation générique de la méthode de nudging sur un pas de temps Δt :

$$\mathbf{X}(t + \Delta t) = \mathbf{X}^*(t + \Delta t) + 2\Delta t \frac{\mathbf{X}^{\text{ref}}(t + \Delta t) - \mathbf{X}^*(t + \Delta t)}{\tau} \quad (4.2)$$

\mathbf{X}^* est le vecteur de la variable pronostique avant le nudging tandis que τ désigne le temps caractéristique du rappel vers les données de référence \mathbf{X}^{ref} . La comparaison des équations 4.1 et 4.2, sachant que $\mathbf{X}^*(t + \Delta t) = \mathbf{X}(t) + \mathbf{M}(\mathbf{X}(t), t)$, montre que la perturbation stochastique additive doit se rapprocher de ce que serait le terme de nudging au temps $t + \Delta t$, afin d'obtenir

une prévision proche de la référence. En cours de prévision, l'état final n'est pas connu *a priori*. Pour chaque membre, un incrément de nudging différent, correspondant à un rappel du modèle à un autre instant t' de la période étudiée, est tiré de manière aléatoire. L'idée est de trouver ainsi une correction d'erreur de tendance initiale du modèle allant dans le sens d'une amélioration de la prévision, en rapprochant les valeurs de \mathbf{X} de celles des données de référence. L'aspect haute fréquence des perturbations doit permettre d'obtenir une dispersion d'ensemble satisfaisante. Le profil vertical de coefficients d'atténuation du nudging permet de laisser la surface libre.

Un choix primordial pour définir la dynamique stochastique est donc celui de la méthode et des données de référence servant au nudging. Deux solutions ont été envisagées dans le cadre de cette thèse :

- garder le modèle CNRM-CM5.1 à l'équilibre et guider les anomalies de celui-ci vers celles des réanalyses ERA-Interim (nudging d'anomalie, NA)
- guider le modèle en plusieurs étapes vers ERA-Interim (nudging itéré, NI)

Les détails d'implémentation de ces méthodes ainsi que les résultats obtenus sont présentés dans la suite de ce chapitre.

4.2 Perturbations dérivées par nudging d'anomalie

Cette partie présente les résultats obtenus en utilisant le nudging d'anomalie pour calculer les populations de perturbations stochastiques $\delta\mathbf{X}$.

4.2.1 Motivations et mise en œuvre

4.2.1.1 Définition des termes $\delta\mathbf{X}$: intégration en nudging d'anomalie

Le choix d'utiliser le nudging d'anomalie pour la méthode de dynamique stochastique est fortement lié au modèle utilisé dans le cadre de cette thèse. Les travaux réalisés dans le cadre du cinquième rapport du GIEC ont permis de définir un ensemble de choix de modélisation et de paramétrisations tel que le modèle couplé CNRM-CM5.1 reste à l'équilibre au cours des simulations climatiques. Dès lors, rappeler la dynamique de l'atmosphère du modèle ARPEGE-Climat vers celle des réanalyses ERA-Interim viendrait perturber l'équilibre du modèle, en le forçant par exemple à se refroidir si la climatologie de la température du modèle à l'équilibre est plus chaude que celle d'ERA-Interim. Les termes de rappel d'une simulation en nudging simple comporteraient alors une majorité de termes de correction refroidissant le modèle. Ces termes ne correspondraient pas nécessairement aux erreurs que fait le modèle quand il est libre d'évoluer, ce qui est le cas lorsqu'il est utilisé en prévision saisonnière. Plutôt que de rappeler directement les champs de température, humidité spécifique et fonction de courant vers ceux d'ERA-Interim, nous rappelons donc ici les anomalies de ces champs par rapport à leur climatologie calculée lors d'une première intégration libre du modèle sur la période d'étude vers les anomalies des champs des réanalyses ERA-Interim par rapport à la climatologie ERA-Interim.

L'équation 4.2 s'écrit alors, en notant $\overline{\mathbf{X}}$ la climatologie de $\mathbf{X}(t)$:

$$\mathbf{X}(t + \Delta t) = \mathbf{X}^*(t + \Delta t) + 2\Delta t \frac{(\mathbf{X}^{\text{ref}}(t + \Delta t) - \overline{\mathbf{X}^{\text{ref}}}) - (\mathbf{X}^*(t + \Delta t) - \overline{\mathbf{X}})}{\tau}$$

Variable	τ
Q	≈ 1 jour
T	≈ 1 jour
Ψ	≈ 5 heures
Niveaux (altitude standard, m)	λ
33	0.
148	0.1
352	0.2
629	0.5
965 m et plus haut	1.

TABLEAU 4.1 – Valeurs de τ et λ pour les expériences en nudging d'anomalie.

En pratique au cours de l'intégration du modèle en mode nudging d'anomalie on calcule pour un jour donné la climatologie du modèle à l'aide d'une interpolation linéaire. Pour la première moitié du mois, la climatologie est interpolée linéairement entre les climatologies mensuelles du mois précédent et du mois courant. Pour la fin du mois, elle est interpolée entre le mois courant et le mois suivant.

Les temps caractéristiques de rappel (τ) et les termes d'atténuation du nudging sur les niveaux les plus proches de la surface (λ) sont donnés dans le tableau 4.1. Ces temps caractéristiques sont relativement courts et imposent donc à l'atmosphère du modèle couplé de suivre les anomalies des réanalyses ERA-Interim. La vorticit  est plus fortement guid e que la temp rature et l'humidit  sp cifique, afin de contraindre les anomalies de circulation atmosph rique   suivre celles des r analyses.

Pour l'int gration du mod le coupl  avec nudging d'anomalie, l'oc an est  galement guid , en rappelant la temp rature de surface de l'oc an avec un coefficient de r troaction dQ/dT pris  gal   $-231.5 \text{ W/m}^2/\text{K}$, comme pr conis  en pr vision saisonni re depuis le projet DEMETER. Les autres composantes de CNRM-CM5.1 ne sont pas rappel es. Le mod le a  t  int gr  ainsi sur 22 ann es   l'aide des donn es ERA-Interim de janvier 1989   f vrier 2011.

Au cours de l'int gration, les termes de rappel pour l'atmosph re sont enregistr s toutes les 6 heures. Ce sont ces termes qui constitueront la base des ensembles $\delta\mathbf{X}$ dans lesquels sont tir es les perturbations stochastiques.

Les fichiers d' tats initiaux des diff rentes composantes, enregistr s chaque d but de mois au cours de cette int gration de CNRM-CM5.1, servent d' tats initiaux des pr visions saisonni res  tudi es dans cette partie.

4.2.1.2 Mise en  uvre d'une r tro-pr vision saisonni re

L'int gration avec nudging d'anomalie sert   d finir la population des perturbations $\{\delta\mathbf{X}\}$ des saisons MJJA et NDJF, ainsi que les  tats initiaux des pr visions saisonni res pour les mois de mai et novembre 1989   2010. Pour chaque saison  tudi e, deux jeux de r tro-pr visions de 15 membres ont  t  impl ment s, avec et sans perturbations stochastiques.

En l'absence de fichiers de conditions initiales quotidiennes pour chacune des composantes du mod le coupl , les ensembles sont initialis s en perturbant pour chaque membre les condi-

tions initiales de la composante ARPEGE-Climat du début du premier mois de prévision par un terme $\delta\mathbf{X}$ aléatoire, ajouté uniquement à l'instant $t = 0$.

Les rétro-prévisions saisonnières avec perturbations initiales sont notées INI, celles avec dynamique stochastique DS. Les prévisions de l'ensemble DS sont perturbées en ajoutant à chaque pas de temps un terme $\delta\mathbf{X}$ tiré aléatoirement parmi $\{\delta\mathbf{X}\}_M$, où M désigne le mois calendaire en cours de prévision. Le même $\delta\mathbf{X}$ est appliqué pendant une période \tilde{t} de 6 heures.

4.2.2 Caractéristiques des perturbations

4.2.2.1 Décomposition des termes de perturbation

Par construction, la moyenne des perturbations d'un champ sur la période de référence devrait être faible par rapport à son écart-type, si le modèle nudgé se comporte de manière similaire au modèle libre. Afin d'étudier la part du terme moyen et de l'écart-type dans les perturbations, la moyenne des carrés des N perturbations se décompose en moyenne et variance :

$$\frac{1}{N} \sum_{i=1}^N \delta X_i^2 = \overline{\delta X}^2 + \text{Var}(\delta X)$$

En classant pour un mois donné les perturbations d'une variable et d'un niveau du modèle δX par année y , la variance des δX de ce mois peut se décomposer en plusieurs termes comme suit :

$$\text{Var}(\delta X) = \frac{1}{N} \sum_{i=1}^N \left(\delta X_i - \overline{\delta X} \right)^2 = \frac{1}{N} \sum_y \sum_{i_y=1}^{n_y} \left(\delta X_{i_y}^{(y)} - \overline{\delta X}^{(y)} \right)^2 + \sum_y \frac{n_y}{N} \left(\overline{\delta X}^{(y)} - \overline{\delta X} \right)^2 \quad (4.3)$$

où n_y désigne le nombre de δX pour l'année y .

Le premier terme du membre de droite est la variance intra-annuelle des corrections, tandis que le deuxième terme désigne la variance interannuelle.

L'équation 4.3 permet de comparer trois termes du carré des perturbations. Formellement :

$$\overline{\delta X^2} = \overline{\delta X}^2 + \text{Var}_{\text{inter}(y)}(\delta X) + \text{Var}_{\text{intra}(y)}(\delta X) \quad (4.4)$$

Dans ce qui suit, les parts de δX^2 expliquées par le terme moyen, la variabilité interannuelle et intra-annuelle sont montrées en calculant les rapports

$$\frac{\overline{\delta X}^2}{\overline{\delta X^2}}, \frac{\text{Var}_{\text{inter}(y)}(\delta X)}{\overline{\delta X^2}} \text{ et } \frac{\text{Var}_{\text{intra}(y)}(\delta X)}{\overline{\delta X^2}}.$$

4.2.2.2 Résultats et interprétation

La figure 4.1 montre la décomposition détaillée au paragraphe précédent pour les saisons d'hiver (NDJF) et d'été (MJJA) sur la période 1989–2010 pour différents niveaux et variables

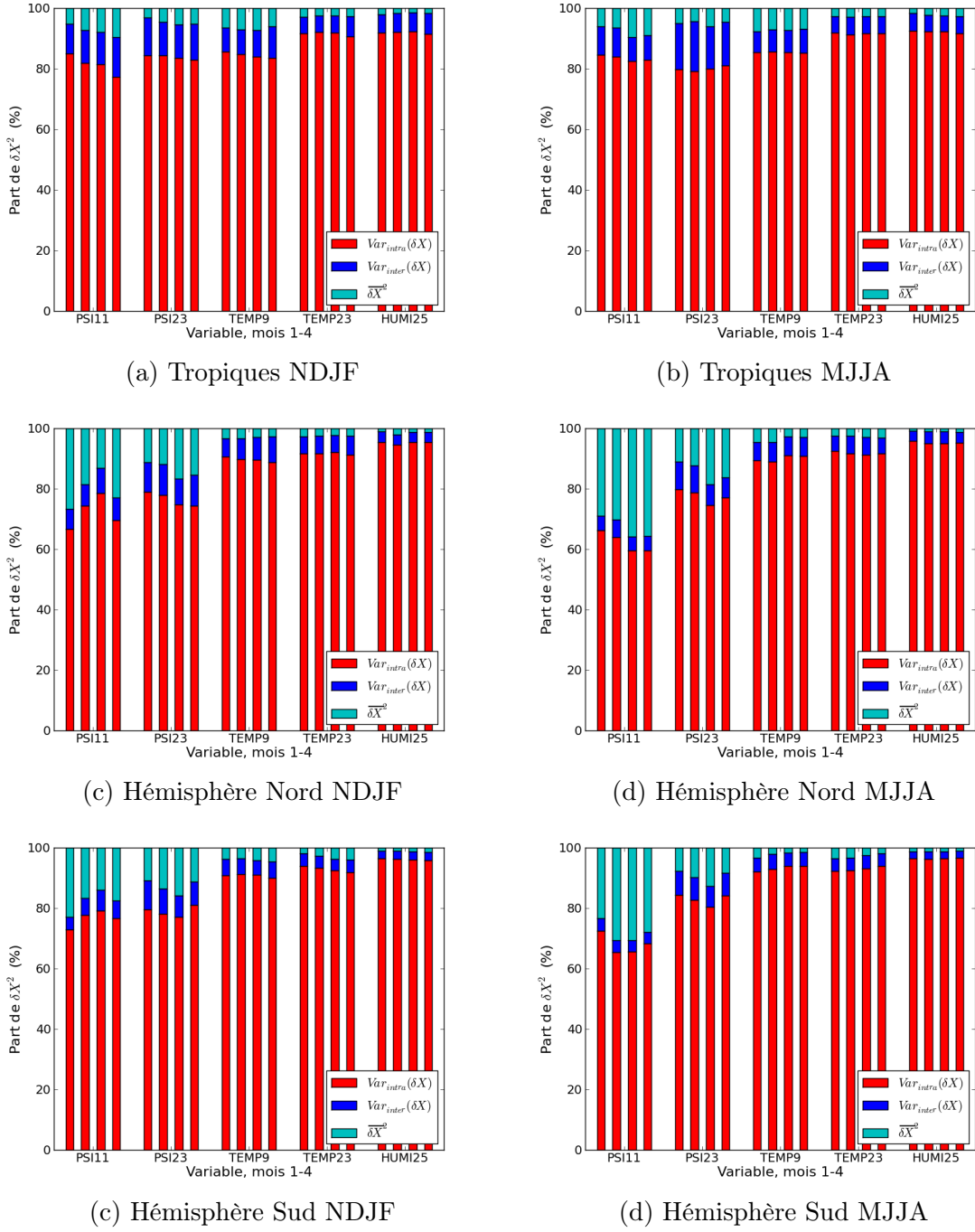


FIGURE 4.1 – Décomposition moyenne des termes quadratiques des perturbations pour NDJF (colonne de gauche) et MJJA (colonne de droite) sur les tropiques (30°N, 30°S) et les zones tempérées des hémisphères Nord et Sud (30° à 75° de latitude) (de haut en bas). Chaque graphique montre le pourcentage de chaque terme de l'équation 4.4 pour les mois 1 à 4 de la saison dans l'intégration de référence nudgée et 5 variables différentes du modèle.

Niveau	Altitude standard	Pression standard
9	12.4 km	181 hPa
10	11.5 km	208 hPa
11	10.7 km	238 hPa
18	5.6 km	500 hPa
23	2.7 km	726 hPa
25	1.8 km	817 hPa

TABLEAU 4.2 – Altitude et pression de l’atmosphère standard pour les niveaux du modèle ARPEGE-Climat évoqués dans les chapitres 4 et 5.

Paramètre	Ensemble Région	INI			DS		
		5%	95%	mACC	5%	95%	mACC
T2m	Niño 3.4	0.921	0.927	0.924	0.927	0.934	0.931
T2m	Tropiques	0.478	0.495	0.487	0.476	0.495	0.486
T2m	HN (30°N–75°N)	0.179	0.217	0.199	0.128	0.216	0.171
Précipitations	Tropiques	0.532	0.547	0.539	0.535	0.546	0.541
Z500	HN (30°N–75°N)	0.215	0.318	0.265	0.112	0.241	0.178

TABLEAU 4.3 – Scores de mACC pour DJF 1989–2010 des ensembles INI et DS. On montre la moyenne des mACC de tirages de 9 membres parmi 15, et les fourchettes de 5% à 95% ainsi obtenues.

pronostiques du modèle. Les altitudes et pressions standard de ces niveaux sont listées dans le tableau 4.2.

Les termes de perturbation moyenne et leur variance interannuelle expliquent en tout entre 10 et 20% de $\overline{\delta X^2}$ sur les tropiques. Sur les latitudes tempérées, les perturbations moyennes sont plus importantes pour la fonction de courant, et atteignent jusqu’à 30% de $\overline{\delta X^2}$ au niveau 11 du modèle (environ 240 hPa) au cours de la saison d’été. Hormis pour la fonction de courant, la variance intra-annuelle constitue plus de 90% de la décomposition de la moyenne des corrections sur les latitudes tempérées. Ceci sous-tend le fait que les perturbations ainsi conçues comportent principalement du bruit pouvant augmenter la dispersion des prévisions du modèle. Il y a peu de différences entre les deux hémisphères, les saisons NDJF et MJJA et les différents mois de ces saisons dans les parts relatives des différentes composantes.

4.2.3 Simulations de la saison d’hiver (DJF)

4.2.3.1 Prévisions déterministes : impact sur la moyenne d’ensemble

Le tableau 4.3 contient les scores moyens de mACC et les fourchettes de 5% à 95% obtenus avec des tirages de 9 membres parmi les 15 des ensembles INI et DS sur plusieurs régions du globe pour la température à 2 mètres (T2m), les précipitations et le géopotentiel à 500 hPa (Z500). Les scores de mACC de la température de surface sur la région Niño 3.4 sont améliorés de manière significative avec les perturbations stochastiques. Par contre, s’il n’y a que très peu de différences entre les ensembles INI et DS sur les tropiques, les scores de mACC de Z500 et de T2m sur l’hémisphère Nord extra-tropical sont significativement dégradés pour l’ensemble DS.

Nous avons vu que la décomposition des perturbations comportait un terme moyen non

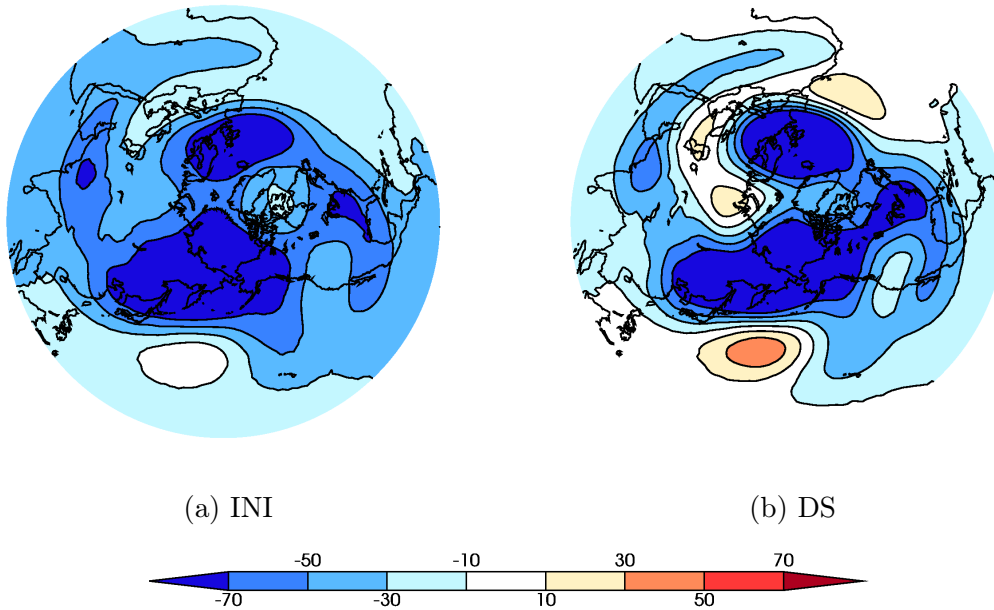


FIGURE 4.2 – Biais (en m) sur l’hémisphère Nord du Z500 de la saison DJF des ensembles INI et DS.

négligeable, en particulier pour la fonction de courant. Ce terme moyen est dû à la tendance qu’a le modèle couplé à s’écarter de sa propre climatologie lorsqu’il est nudgé en mode anomalie, et ne correspond pas aux erreurs que va faire le modèle par rapport aux données de référence lorsqu’il est libre en cours de prévision. Dans le cas où ces deux valeurs ne sont pas de même signe, la perturbation du modèle par les termes $\delta\mathbf{X}$ éloignera encore le modèle de la réalité (dans notre cas, des réanalyses ERA-Interim considérées comme telles). La figure 4.2 montre le biais en DJF de Z500 vis-à-vis d’ERA-Interim des ensembles REF et INI sur la période d’étude. Le modèle couplé a une hauteur de géopotential plus basse qu’ERA-Interim sur l’hémisphère Nord pendant l’hiver, et ces biais moyens sont modifiés par l’introduction de perturbations stochastiques issues du nudging d’anomalie. En particulier, les gradients de biais sont amplifiés sur le Pacifique Nord et sur l’Atlantique, ce qui pourrait avoir une influence sur la circulation atmosphérique sur ces régions.

4.2.3.2 Prévisions probabilistes

Les scores de RPSS et la décomposition en fiabilité et résolution sont donnés dans le tableau 4.4 pour plusieurs paramètres pour DJF 1989–2010. La différence entre les scores sur la région Niño 3.4 et les autres zones d’étude est frappante. Sur cette région, la résolution des ensembles atteint presque la valeur de l’incertitude. Nous retrouvons généralement les mêmes différences de scores entre les ensembles INI et DS que pour le mACC. Les scores sont équivalents sur les tropiques et meilleurs pour l’ENSO, mais moins bons sur l’hémisphère Nord (et même nettement dégradés dans le cas du Z500). Il est intéressant de nuancer les résultats en termes de RPSS par les composantes de fiabilité et résolution : si pour la T2m sur la région Niño 3.4 les deux composantes sont améliorées par la dynamique stochastique, pour les précipitations sur les tropiques, la baisse du score de RPSS est due principalement à une dégradation de

Paramètre	Ensemble Région	INI			DS		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.201	0.959	0.758	0.202	0.970	0.768
Précipitations	Tropiques	0.557	0.597	0.041	0.555	0.590	0.035
Z500	HN	0.479	0.512	0.033	0.505	0.506	0.001
T2m	HN	0.543	0.546	0.003	0.539	0.537	-0.002

TABEAU 4.4 – Scores de RPSS et décomposition en termes de fiabilité (Rel, divisé ici par l’incertitude Unc) et résolution (Res) pour les ensembles INI (à gauche) et DS (à droite) de prévisions saisonnières de DJF 1989–2010.

la résolution de l’ensemble. Pour le Z500 sur l’hémisphère Nord, les deux composantes sont détériorées avec la dynamique stochastique, tandis que pour la T2m la fiabilité est légèrement améliorée mais cette amélioration est compensée par une baisse de la résolution.

Notons que le terme de fiabilité est lié à la calibration du modèle, et peut donc être amélioré *a posteriori*. Il est plus difficile d’obtenir des améliorations de la résolution. L’augmentation de la résolution sur la région Niño 3.4 est donc un apport incontestable de la méthode sur cette région, alors qu’elle échoue sur les autres régions étudiées.

4.2.3.3 Impact sur la variabilité intra-saisonnière

Un point fondamental à explorer est celui de l’impact des perturbations stochastiques sur les caractéristiques du modèle à l’échelle intra-saisonnière. A trop vouloir perturber la dynamique du modèle, l’a-t-on secoué au point de lui faire créer des états irréalistes et trop instables ? Les travaux de Jung et al. (2005) ont par exemple montré que la méthode SPPT avait peu d’influence sur les régimes de temps sur le Pacifique Nord simulés par le modèle de prévision à moyen terme du CEPMMT, mais que la méthode CASBS permettait d’améliorer les fréquences de ces régimes.

On s’intéresse ici à l’impact de la dynamique stochastique sur les régimes de temps de la région Atlantique Nord – Europe, définie par les latitudes de 20 à 80°N et les longitudes de 90°W à 30°E. Les régimes de temps hivernaux (DJF) sur cette région sont calculés suivant la méthode suggérée par Michelangeli et al. (1995) sur la période des réanalyses ERA-Interim (1979–2010) en utilisant les séries quotidiennes du champ de géopotential à 500 hPa (voir la partie 2.3.3 pour plus de détails). Les différents ensembles sont comparés en projetant les séries de Z500 corrigées de leur biais moyen sur la base des EOF déterminée à partir des réanalyses ERA-Interim, puis en calculant la distance euclidienne de ces projections à chacun des quatre régimes. Un seuil de 3 jours est fixé de manière à éliminer les champs de Z500 n’appartenant pas à un régime particulier, correspondant à une transition entre deux types de temps. Cette méthode permet de calculer les fréquences et durées moyennes prévues pour chaque régime et chaque saison DJF de la période de rétro-prévision.

Le tableau 4.5 recense les durées moyennes et les fréquences de chaque régime pour DJF 1989–2010 dans les réanalyses ERA-Interim et les ensembles INI et DS. La durée de chaque régime est réduite avec l’introduction de la dynamique stochastique, alors que les durées moyennes de l’ensemble INI étaient assez proches des réanalyses. Ceci montre que l’introduction des perturbations va pousser le modèle à explorer plus rapidement l’espace des phases en passant plus rapidement d’un régime à un autre. Ce point est également confirmé par le

Régime	Durée (jours)			Fréquence			
	ERA-I	INI	DS	ERA-I	INI	DS	
DA	6.17	6.35	5.72	0.16	0.17	<i>0.021</i>	0.17 <i>0.016</i>
Zonal	9.68	8.03	6.88	0.32	0.29	<i>0.019</i>	0.29 <i>0.032</i>
Blocage	7.65	7.22	6.74	0.25	0.23	<i>0.012</i>	0.21 <i>0.018</i>
NAO ⁻	8.51	8.76	8.24	0.18	0.20	<i>0.033</i>	0.19 <i>0.032</i>

TABLEAU 4.5 – Durée (en jours) et fréquence moyenne (en %) de chaque régime pour les réanalyses ERA-Interim et les ensembles INI et DS (moyenne d'ensemble et écart-type en italique) pour DJF 1989–2010.

nombre de jours non classés (régimes de durée inférieure à 3 jours) qui est quasiment doublé par l'introduction de la dynamique stochastique. En matière de fréquences d'occurrence des régimes, peu de changement entre INI et DS est visible. Tous deux reproduisent des fréquences climatologiques proches des réanalyses. Seule la fréquence du régime de blocage est légèrement dégradée par l'introduction de la dynamique stochastique. L'écart-type des membres est légèrement diminué pour le régime Dorsale Atlantique, tandis qu'il augmente pour les régimes Zonal et Blocage.

Les scores de corrélation des fréquences saisonnières de chacun des régimes sont donnés dans le tableau 4.6. La dégradation des prévisions avec l'introduction de la dynamique stochastique est alors évidente, puisque tous les scores sont moins bons que ceux avec l'ensemble INI. Au vu des résultats obtenus en termes de mACC pour Z500, cette conclusion n'est pas surprenante. On peut également calculer un score probabiliste comparant la distance entre fréquences prévues et fréquences observées pour les quatre régimes dans les prévisions et les données des réanalyses. Pour les prévisions, cette distance s'écrit

$$BS_{mod} = \sum_{i_r=1}^4 \frac{1}{N} \sum_{y=1}^N (f_{mod}^y(i_r) - f_{obs}^y(i_r))^2$$

tandis que pour la climatologie on écrit

$$BS_{clim} = \sum_{i_r=1}^4 \frac{1}{N} \sum_{y=1}^N \left(\overline{f_{obs}^{y \neq y}}(i_r) - f_{obs}^y(i_r) \right)^2$$

Ces distances s'apparentent à des scores de Brier pour les régimes. Un Brier Skill Score ou BSS est défini en calculant

$$BSS = 1 - \frac{BS_{mod}}{BS_{clim}}$$

Dans la dernière colonne du tableau 4.6 on note que ce score reste inchangé par l'introduction de la dynamique stochastique.

Régime	DA	Zonal	Blocage	NAO ⁻	BSS
INI	0.26	0.19	0.39	0.10	0.21
DS	0.12	0.13	0.03	0.04	0.19

TABLEAU 4.6 – Corrélation des fréquences moyennes prévues par INI et DS avec les fréquences des réanalyses pour chacun des régimes, et score de BSS (voir texte).

Paramètre	Ensemble Région	INI			DS		
		5%	95%	mACC	5%	95%	mACC
T2m	Niño 3.4	0.726	0.756	0.741	0.734	0.763	0.750
T2m	Tropiques	0.367	0.379	0.373	0.384	0.399	0.392
T2m	HN (30°N–75°N)	0.166	0.198	0.182	0.148	0.185	0.167
Précipitations	Tropiques	0.355	0.371	0.363	0.375	0.390	0.382
Z500	HN (30°N–75°N)	0.104	0.188	0.147	0.018	0.084	0.050
Z500	HS (30°S–75°S)	0.130	0.216	0.173	0.070	0.168	0.119

TABLEAU 4.7 – Idem tableau 4.3, pour la saison JJA.

4.2.4 Simulations de la saison d’été (JJA)

Le tableau 4.7 regroupe les scores moyens et les fourchettes de mACC pour 9 membres parmi les 15 des ensembles INI et DS en JJA. On retrouve des résultats très similaires à ceux de la saison DJF, même si les scores sont généralement moins bons en été (en particulier pour les précipitations sur les tropiques). Les scores sur les tropiques sont significativement améliorés par les perturbations stochastiques. Par contre, il y a toujours une nette dégradation des scores de Z500 sur les latitudes tempérées boréales et australes. Sur la région Niño 3.4, les scores sont nettement moins élevés qu’en DJF, et l’amélioration des corrélations d’anomalie avec l’ensemble DS n’est plus significative.

Comme pour la saison DJF, on peut étudier l’influence de la dynamique stochastique sur le RPSS et les termes de fiabilité et de résolution en JJA, montrés dans le tableau 4.8. Ici encore, le RPSS augmente sur la région Niño 3.4 pour la T2m, avec une nette amélioration des deux composantes du score. On note une légère amélioration du RPSS pour les précipitations sur les tropiques, même si la résolution est dégradée. Sur l’hémisphère Nord, les scores sont nettement dégradés avec DS pour le Z500 et diminuent légèrement pour la T2m. Contrairement à la saison d’hiver, ces scores sont moins bons en raison d’une dégradation de la fiabilité et non pas une perte de résolution.

Paramètre	Ensemble Région	INI			DS		
		Rel/U	Res/U	RPSS	Rel/U	Res/U	RPSS
T2m	Niño 3.4	0.516	0.753	0.237	0.507	0.770	0.263
Précipitations	Tropiques	0.565	0.586	0.021	0.556	0.580	0.024
Z500	HN	0.419	0.431	0.012	0.436	0.432	-0.004
T2m	HN	0.454	0.454	0.000	0.456	0.454	-0.002

TABLEAU 4.8 – Idem tableau 4.4, pour JJA 1989–2010.

4.2.5 Discussion

Les résultats obtenus dans cette partie sont surprenants à bien des égards. Pour commencer, la méthode de nudging d'anomalie devait permettre d'obtenir des perturbations de moyenne quasi-nulle, or la figure 4.1 montre bien que ce n'est pas le cas. Le comportement du modèle une fois nudgé est visiblement bien différent de celui du modèle en cours de prévision (libre). Les perturbations ajoutées ne s'apparentent donc pas à du bruit blanc, mais à des termes comportant une composante moyenne de rappel qui ne correspond pas nécessairement aux erreurs que fait vraiment le modèle. Ceci peut avoir pour conséquence d'accentuer les biais de certaines variables et de dégrader les prévisions. C'est peut-être ce qui est à l'origine de la dégradation des prévisions de Z500 et de température sur les régions tempérées avec cette méthode.

Une fois fait ce constat, on peut se demander comment développer une méthode qui porterait une composante moyenne non nulle, mais qui corrigerait les écarts du modèle aux données de référence. L'idée du nudging d'anomalie était de garder le modèle dans son équilibre, on peut au contraire inciter le modèle à se rapprocher des réanalyses. C'est l'objet de la méthode de nudging itéré qui est décrite dans la partie suivante.

4.3 Perturbations dérivées par nudging itéré

Dans cette partie nous verrons les résultats obtenus en utilisant la méthode de nudging itéré pour estimer la population des perturbations $\{\delta\mathbf{X}\}$.

4.3.1 Motivations et mise en œuvre

Cette méthode est moins directe à mettre en œuvre que la première solution. En guidant directement l'atmosphère du modèle couplé vers les réanalyses ERA-Interim, il est difficile d'isoler les erreurs de tendance initiale du modèle des rappels liées aux écarts de la climatologie du modèle à celle d'ERA-Interim. Utiliser ces termes de rappel en cours de prévision pourrait alors avoir l'effet inverse de celui recherché, si par exemple le modèle couplé est plus froid qu'ERA-Interim à l'équilibre mais a des erreurs de tendance initiale chaudes pendant la phase de spin-up. Dès lors, pourquoi utiliser le nudging simple, alors même qu'on dispose dans le cadre de cette thèse d'un modèle couplé à l'équilibre ? La motivation principale de cette partie est de développer une méthode plus générique, qui pourra s'appliquer à des variantes du modèle couplé qui peuvent dériver aux échelles climatiques. La méthode développée ici a notamment pu être testée en marge de la thèse avec différentes résolutions horizontales et verticales du modèle ARPEGE-Climat.

4.3.1.1 Définition des termes $\delta\mathbf{X}$

Contrairement à la méthode avec nudging d'anomalie, les termes de rappel qui constituent les ensembles $\{\delta\mathbf{X}\}$ ne proviennent pas directement d'une première intégration du modèle nudgé vers la référence. La première intégration est nudgée directement vers ERA-Interim suivant l'équation 4.2. Les mêmes caractéristiques (τ, λ) que pour le nudging d'anomalie sont

Variable	τ
Q	≈ 1 mois
T	≈ 1 mois
Ψ	≈ 10 jours
Niveaux (m)	λ
33	0.
148	0.01
352	0.03
629	0.1
965	0.3
1350 m et plus haut	1.

TABLEAU 4.9 – Valeurs de τ et λ pour la deuxième étape du nudging itéré (nudging faible).

utilisées (voir le tableau 4.1, page 113). L’océan est également nudgé en surface et en profondeur vers les réanalyses NEMOVAR, sauf dans les tropiques. Cette intégration couvre la période janvier 1979 à décembre 2010, soit 10 ans de plus que la méthode de nudging d’anomalie.

Dans un deuxième temps, afin de représenter de manière plus réaliste les erreurs de tendance initiale du modèle en cours de prévision, un ensemble de rétro-prévisions saisonnières nudgées est implémenté. Pour chaque mois d’initialisation des prévisions, quatre mois de simulation sont calculés à partir des états initiaux obtenus lors de la première intégration du modèle. Contrairement à une prévision saisonnière libre, cette intégration est rappelée faiblement vers la première intégration avec la technique de nudging simple. Les temps caractéristiques et les coefficients d’atténuation vers la surface de ce nudging faible sont listés dans le tableau 4.9. Quatre membres sont calculés pour chaque saison, à l’aide de perturbations initiales tirées parmi les rappels de la première intégration. Pour chaque membre, les termes de rappel faible sont sauvegardés quotidiennement. On obtient ainsi une population de perturbations stochastiques $\{\delta\mathbf{X}\}$ provenant des quatre membres de la simulation nudgée faiblement.

4.3.1.2 Caractéristiques des termes de correction

Contrairement à la méthode de nudging d’anomalie, les termes de rappel n’ont aucune raison d’être centrés. En particulier, au vu de la méthode utilisée, ces rappels peuvent dépendre du délai de la prévision. Nous allons maintenant regarder les moyennes mensuelles des perturbations au cours des rétro-prévisions nudgées d’été et d’hiver, pour les trois variables nudgées et différents niveaux du modèle, afin d’évaluer l’impact de ce délai de prévision sur l’évolution des perturbations.

La figure 4.3 (page 124) montre les corrections moyennes pour les mois de mai, juin, juillet et août des intégrations en nudging itéré pour la saison d’été. Les perturbations sont montrées aux niveaux du modèle où elles sont les plus fortes ou les plus contrastées en moyenne zonale et saisonnière. Pour l’humidité au niveau 25 du modèle (environ 820 hPa dans l’atmosphère standard) il y a peu d’évolution des perturbations d’un mois à l’autre, si ce n’est un changement de signe du dipôle de corrections d’humidité sur l’Atlantique tropical entre les mois de juin et juillet. Pour la fonction de courant au niveau 11 (environ 240 hPa) les corrections moyennes s’intensifient au fur et à mesure du délai de prévision, tout particulièrement sur les tropiques. Quant à la température au niveau 9 (environ 180 hPa), ses corrections semblent s’atténuer

sur la bande tropicale et s'intensifier dans les moyennes latitudes et les régions polaires entre un mois et quatre mois d'échéance. Elles sont toujours positives en moyenne, ce qui indique que le modèle ARPEGE-Climat est trop froid par rapport à ERA-Interim à ce niveau vertical et dans cette configuration.

Les mêmes champs moyens mensuels sont montrés pour la saison NDJF dans la figure 4.4 (page 125). Les évolutions des corrections d'erreurs de tendance initiale sont très similaires à MJJA, même si les corrections sont généralement d'amplitude moindre.

Contrairement à la méthode basée sur le nudging d'anomalie, le terme moyen des perturbations englobe le biais du modèle par rapport à la référence ERA-Interim. On peut décomposer de la même manière que pour le nudging d'anomalie la moyenne des carrés des perturbations, afin de représenter la part des différents termes (voir l'équation 4.4, page 114).

La figure 4.5 montre les décompositions obtenues pour les saisons MJJA et NDJF, pour plusieurs variables et niveaux du modèle, pour les tropiques et les régions tempérées de l'hémisphère Nord et de l'hémisphère Sud.

En comparant la figure 4.5 (a) aux figures (c) et (e), il apparaît que la part du terme de correction quadratique moyenne est plus importante sur les tropiques que sur les régions tempérées. C'est particulièrement frappant pour la température au niveau 9 du modèle où la correction moyenne représente environ 80% de δX^2 . Le même phénomène est visible pour les corrections de la saison d'été (figures (b), (d) et (f)). L'évolution mois par mois de la part des perturbations moyennes dans la décomposition dépend fortement des variables et saisons regardées. Sur les tropiques, la part des corrections moyennes dans la décomposition augmente avec le délai de prévision pour la fonction de courant au niveau 11, la température au niveau 23 (environ 725 hPa) et l'humidité au niveau 25, et évolue peu pour la fonction de courant au niveau 23 et la température au niveau 9, ce pour les deux saisons étudiées. Sur les régions tempérées, la part des corrections moyennes augmente nettement avec le délai de prévision pour la température au niveau 9 en MJJA sur l'hémisphère Nord, et symétriquement en NDJF sur l'hémisphère Sud. Les autres variables ont une évolution moins prononcée.

En confrontant ces résultats à ceux obtenus avec la technique basée sur le nudging d'anomalie (figure 4.1 page 115) il apparaît que mis à part le cas de la fonction de courant, les termes de correction moyenne et de variance interannuelle sont plus importants avec le nudging itéré.

De manière générale, il est intéressant de noter que la variance intra-annuelle, et donc la haute fréquence, prédomine encore dans la décomposition de la plupart des termes de perturbation. Malgré la part de correction moyenne correspondant à une rectification des biais du modèle, les perturbations sont dominées par un terme de haute fréquence.

4.3.1.3 Rétro-prévisions saisonnières

L'utilisation des perturbations en prévision saisonnière se fait de manière similaire à la méthode de nudging d'anomalie, en classant les corrections suivant le mois auquel elles correspondent. Ce choix est d'autant plus important que les corrections sont calculées par une implémentation du modèle en mode prévision saisonnière. Pour un mois M donné, les corrections correspondent ainsi au même délai du modèle par rapport à son initialisation.

Pour les méthodes purement aléatoires, les corrections des quatre membres de l'intégration faiblement nudgée sont regroupées en une seule population $\{\delta \mathbf{X}\}_M$. Sauf mention du contraire,

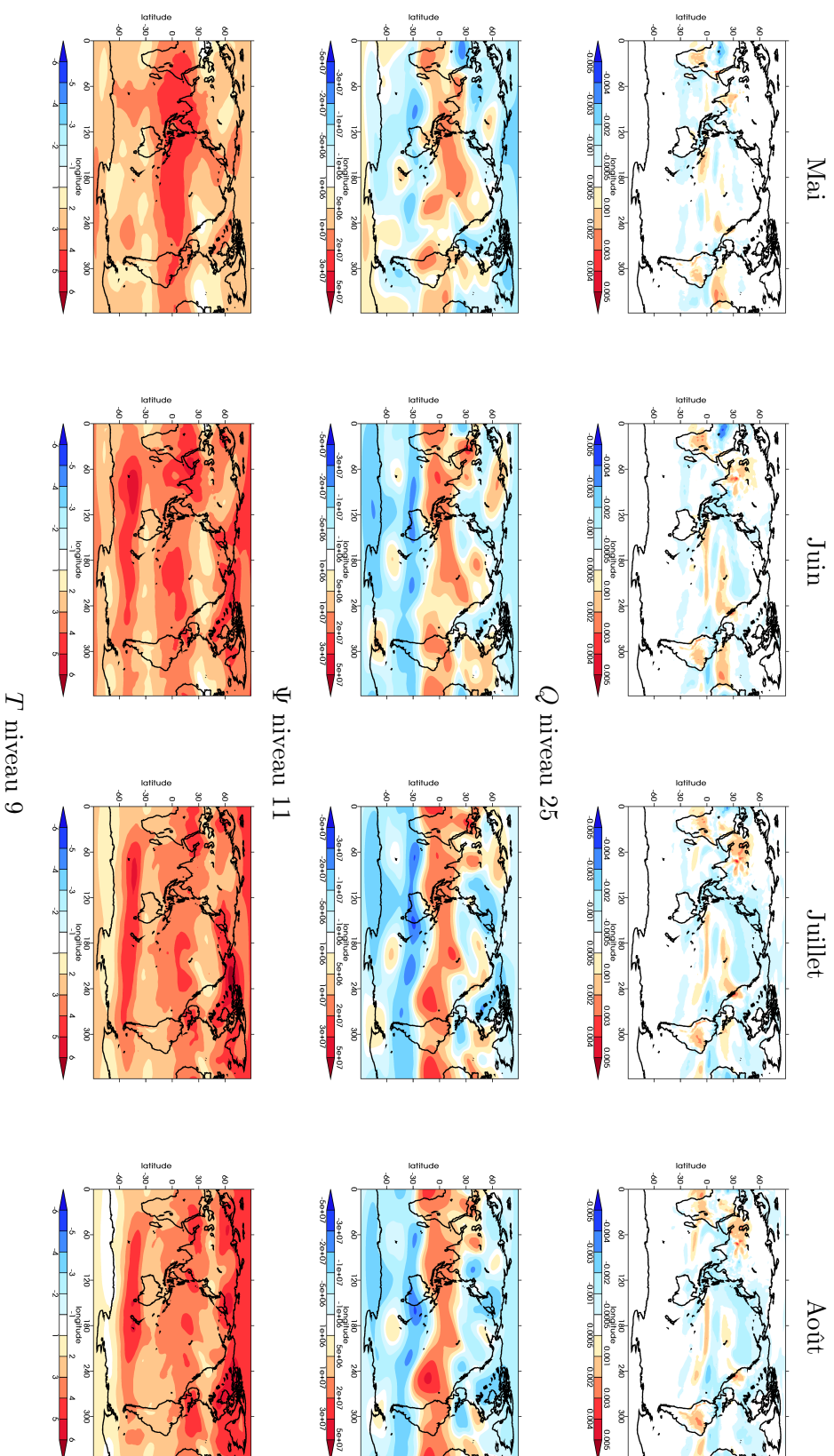


FIGURE 4.3 – Moyennes des perturbations δX pour la prévision nudgée lancée en mai en fonction du délai de prévision (de gauche à droite : mai, juin, juillet et août) pour différents niveaux et variables nudgées (de haut en bas : humidité au niveau 25, fonction de courant au niveau 11, température au niveau 9).

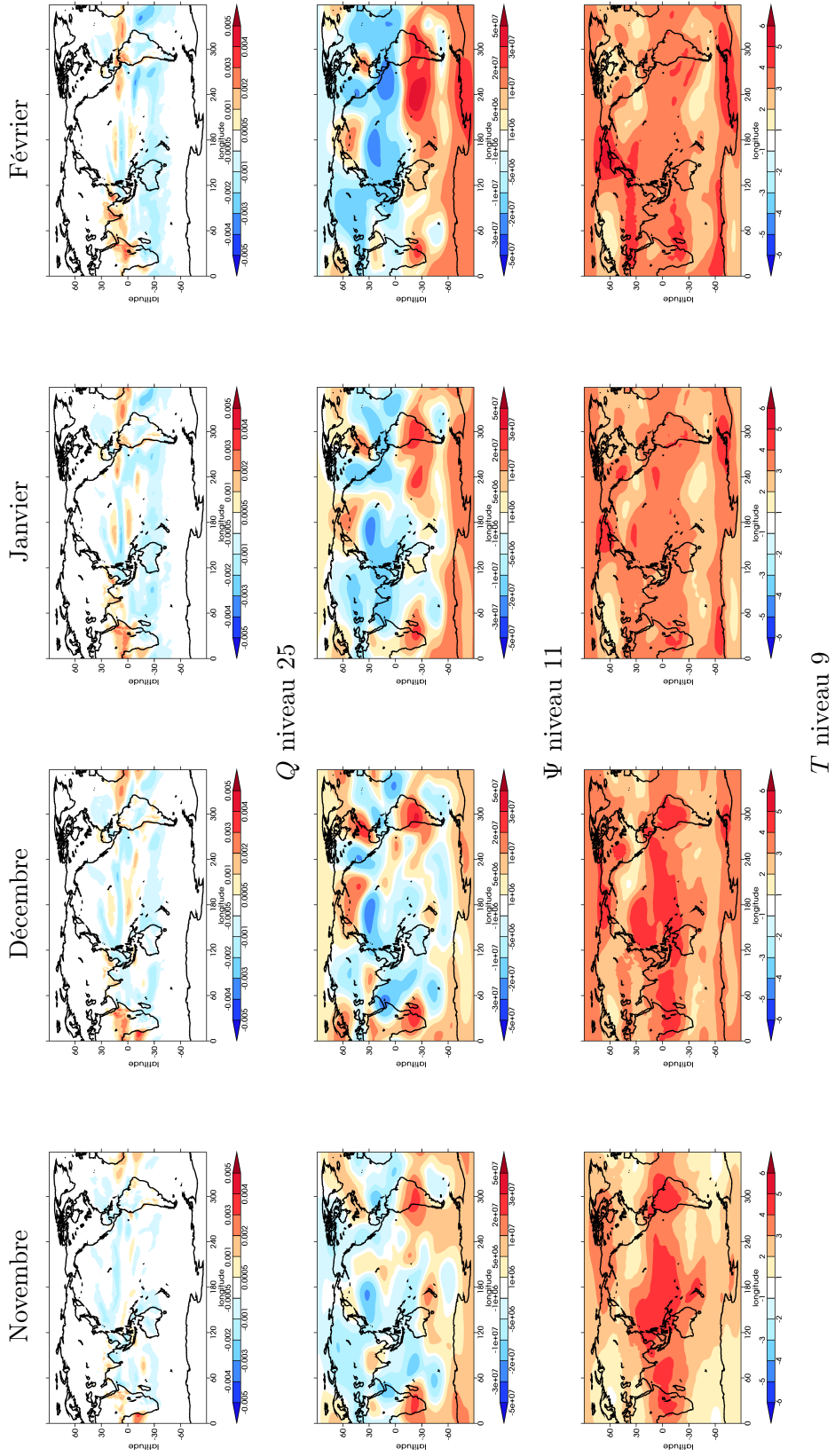


FIGURE 4.4 – Idem figure 4.3, pour la prévision nudgée lancée en novembre pour la saison d’hiver (de gauche à droite : novembre, décembre, janvier et février).

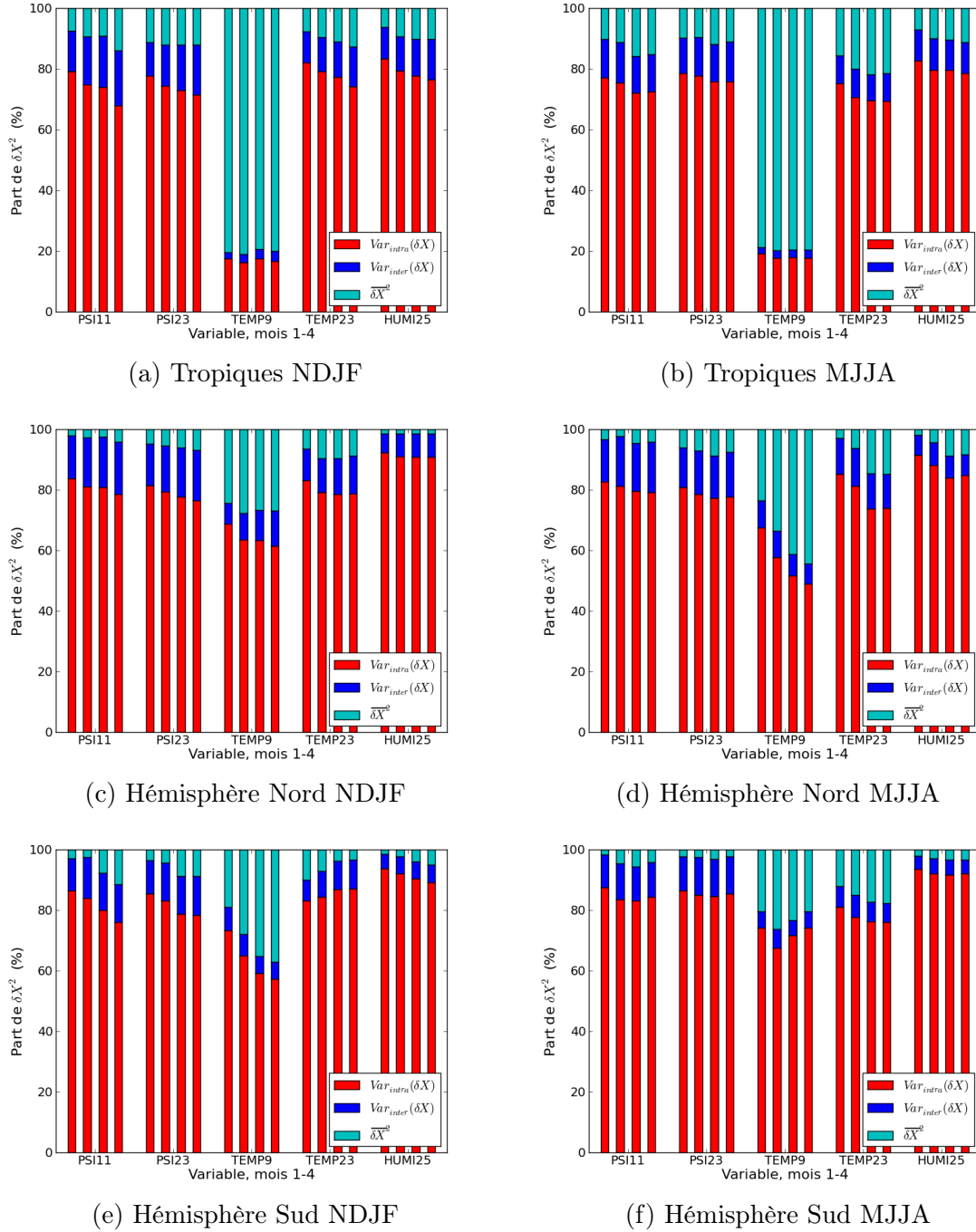


FIGURE 4.5 – Décomposition moyenne des termes quadratiques des perturbations pour NDJF (colonne de gauche) et MJJA (colonne de droite) sur les tropiques (30°N , 30°S) et les zones tempérées des hémisphères Nord et Sud (30° à 75° de latitude) (de haut en bas). Chaque graphique montre le pourcentage de chaque terme de l'équation 4.4 pour les mois 1 à 4 de l'ensemble de quatre prévisions saisonnières nudgées et 5 variables différentes du modèle.

une nouvelle correction est tirée toutes les 6 heures au cours du calcul de la prévision saisonnière.

L'initialisation des prévisions saisonnières se fait à partir des fichiers de conditions initiales de la première intégration du nudging itéré, à l'exception de l'océan initialisé à partir des réanalyses NEMOVAR. Cette décision fait suite au constat que les prévisions initialisées à partir des fichiers de conditions initiales de l'océan de la première intégration donnaient des scores nettement dégradés sur les régions tropicales pour la saison d'été. En particulier, le score de mACC des T2m sur la région Niño 3.4 était inférieur à 0.6, sans doute en raison d'incohérences entre les TSO nudgées et les couches plus profondes de l'océan laissées libres sur les tropiques.

4.3.1.4 Evaluation

Les expériences présentées par la suite portent sur les saisons d'été (JJA initialisée début mai) et d'hiver (DJF initialisée début novembre) pour des prévisions de 2 à 4 mois d'échéance. Afin d'établir les apports et faiblesses des méthodes envisagées, les ensembles avec dynamique stochastique sont comparés à une prévision de référence réalisée dans les mêmes conditions, c'est à dire partant des mêmes états initiaux de l'intégration nudgée, mais avec des perturbations de chaque membre à $t = 0$ uniquement. Cette expérience de 15 membres est notée CM5 INI¹.

4.3.2 Dynamique stochastique aléatoire

4.3.2.1 Définition

La première expérience réalisée avec la méthode de dynamique stochastique basée sur le nudging itéré est une expérience avec tirage aléatoire des perturbations, parmi les corrections correspondant au mois calendaire courant pour les autres années de la période d'étude. Formellement, on tire au cours de la prévision du mois M d'une année Y les $\delta\mathbf{X}$ parmi $\{\delta\mathbf{X}\}_{(M,y \neq Y)}$.

Ces ensembles sont notés DS RAND et comportent 15 membres.

4.3.2.2 Prévisions d'hiver

Les deux premières colonnes de scores du tableau 4.10 correspondent aux mACC des ensembles CM5 INI et DS RAND pour la saison DJF. En prenant en compte les 5^e et 95^e centiles des tirages de 9 membres parmi les 15, on arrive aux conclusions suivantes : le mACC de T2m sur la région Niño 3.4 est dégradé par la dynamique stochastique, tandis que les scores de la T2m et des précipitations sur les tropiques sont améliorés. Sur l'hémisphère Nord, le mACC de Z500 est significativement meilleur avec l'introduction de la dynamique stochastique, et la légère dégradation du score pour la température n'est pas significative.

L'amélioration de la corrélation des anomalies de Z500 sur l'hémisphère Nord se fait conjointement à une nette réduction des biais de Z500 sur la région, comme le montre le haut de la figure 4.6. En particulier, le biais négatif sur l'Arctique est atténué, et le gradient

1. Comme son nom l'indique, c'est une des expériences dont les scores des précipitations sur l'Afrique sont détaillés dans la partie 3.3 du chapitre précédent.

Paramètre	Région	CM5 INI	DS RAND	MMENS	DS MENS	DS SEQ5
T2m	Niño 3.4	0.917	0.911	<i>0.903</i> ^a	0.905	0.912
T2m	Tropiques	0.527	0.534	0.526	0.528	0.526
T2m	HN	0.25	0.23	0.26	<i>0.20</i>	0.25
Précip.	Tropiques	0.553	0.563	<i>0.539</i>	0.553	0.569
Z500	HN	<i>0.21</i> ^b	0.25	0.27	0.21	0.27
Z500	HS	0.28	0.25	0.25	0.25	0.31 ^c

^a CM5 INI, DS RAND, DS SEQ5

^b DS RAND, MMENS, DS SEQ5

^c DS RAND, MMENS, DS MENS

TABLEAU 4.10 – Scores moyens de mACC pour 9 membres parmi 15, pour la saison DJF et différents ensembles avec ou sans dynamique stochastique. Pour les meilleurs et moins bons ensembles, les écarts de scores significatifs avec les autres ensembles sont signalés par des caractères gras et italiques, respectivement. Lorsque ces écarts sont significatifs pour trois des quatre autres ensembles, on précise lesquels.

Paramètre	Ensemble Région	CM5 INI			DS RAND		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.293	0.909	0.616	0.298	0.908	0.610
Précipitations	Tropiques	0.469	0.513	0.044	0.465	0.513	0.048
Z500	HN	0.397	0.416	0.019	0.408	0.436	0.028
T2m	HN	0.461	0.481	0.019	0.465	0.486	0.021

TABLEAU 4.11 – Idem tableau 4.4 mais pour les expériences basées sur le nudging itéré pour DJF 1979–2010.

de biais sur l’Atlantique Nord est considérablement réduit par l’introduction de la dynamique stochastique aléatoire. Aucune amélioration similaire des biais de Z500 n’est constatée sur l’hémisphère Sud pour l’été austral (voir le bas de la figure 4.6).

Afin d’évaluer la qualité des prévisions probabilistes, les RPSS sont calculés pour les terciles de précipitation, température de surface et Z500 sur différentes régions du globe. Les résultats et la décomposition en fiabilité et résolution sont donnés dans le tableau 4.11. Comme pour le mACC, les scores sont améliorés, sauf sur la région Niño 3.4. L’amélioration la plus nette est obtenue pour Z500 sur l’hémisphère Nord. Les différences de RPSS entre les ensembles CM5 INI et DS RAND pour T2m sur la région Niño 3.4 s’expliquent en grande partie par une perte de fiabilité des prévisions. L’augmentation du RPSS des précipitations tropicales est due à une amélioration de la fiabilité, tandis que sur l’hémisphère Nord c’est la résolution qui augmente et la fiabilité qui se dégrade avec la dynamique stochastique.

4.3.2.3 Prévisions d’été

Pour la saison JJA, les scores de mACC sont montrés dans les deux premières colonnes du tableau 4.12. L’impression générale est celle d’une dégradation des scores de mACC avec l’introduction de la dynamique stochastique. La moyenne des tirages de 9 membres parmi 15 de l’ensemble CM5 INI est meilleure que le 95^e centile des scores de DS RAND pour la T2m et les précipitations sur les tropiques, la T2m sur la région Niño 3.4 ainsi que Z500 sur

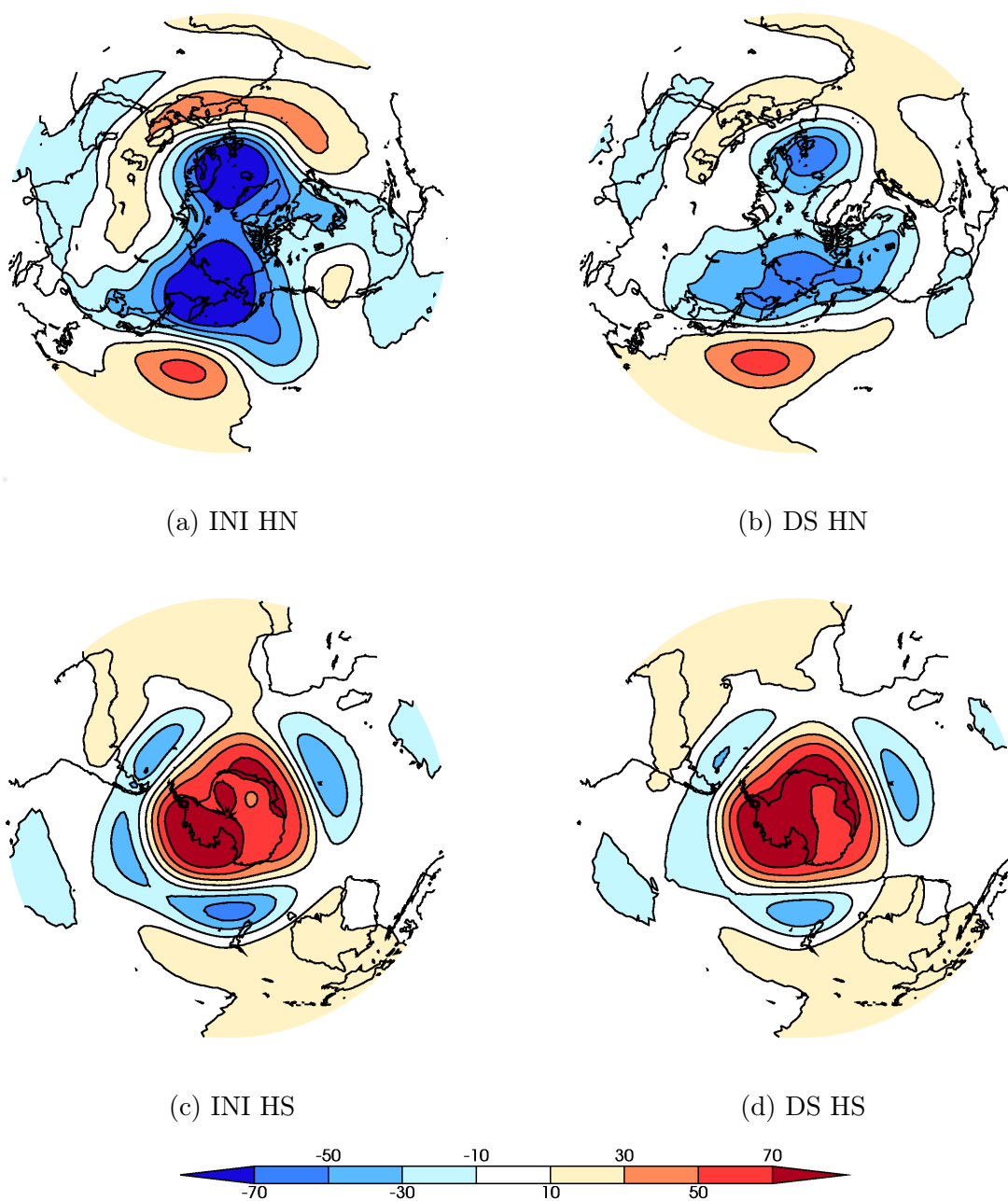


FIGURE 4.6 – Biais (en m) sur l'hémisphère Nord (en haut) et l'hémisphère Sud (en bas) du Z500 de la saison DJF des ensembles CM5 INI et DS RAND.

Paramètre	Région	CM5 INI	DS RAND	MMENS	DS MENS	DS SEQ5
T2m	Niño 3.4	0.845	0.826	0.816	0.808 ^a	0.833
T2m	Tropiques	0.509	0.497	0.497	0.500	0.513^b
T2m	HN	0.204	0.192	0.190	0.218	0.206
Précip.	Tropiques	0.422^b	0.410	0.407	0.406	0.419
Z500	HN	0.14^c	0.10	0.07	0.14	0.09
Z500	HS	0.29^d	0.23	0.27	0.20	0.23

^a CM5 INI, DS RAND, DS SEQ5

^b DS RAND, MMENS, DS MENS

^c DS RAND, MMENS, DS SEQ5

^d DS RAND, DS MENS, DS SEQ5

TABLEAU 4.12 – Idem tableau 4.10, pour la saison JJA.

Paramètre	Ensemble Région	CM5 INI			DS RAND		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.395	0.773	0.378	0.398	0.762	0.363
Précipitations	Tropiques	0.503	0.533	0.030	0.504	0.532	0.028
Z500	HN	0.371	0.380	0.009	0.383	0.385	0.002
T2m	HN	0.404	0.407	0.003	0.404	0.404	0.001

TABLEAU 4.13 – Idem tableau 4.11, pour la saison JJA 1979–2010.

l'hémisphère Sud et l'hémisphère Nord. La dynamique stochastique aléatoire a donc tendance à dégrader les prévisions déterministes de la saison d'été boréal. Qui plus est, contrairement à l'hiver boréal, les scores de mACC pour Z500 en hiver austral sur les latitudes tempérées sont significativement dégradés avec les perturbations aléatoires.

Les scores de RPSS du tableau 4.13 confirment cette tendance, puisqu'ils sont systématiquement moins bons pour l'ensemble DS RAND que pour CM5 INI. La résolution du modèle est dégradée par la dynamique stochastique pour la T2m sur les régions Niño 3.4 et l'hémisphère Nord, tandis que c'est principalement la moins bonne fiabilité qui réduit le RPSS de Z500 sur l'hémisphère Nord.

Une méthode qui donnait de bons résultats pour la saison d'hiver boréal n'est donc pas nécessairement adaptée aux autres saisons sur les latitudes tempérées. La figure 4.7 montre les biais de Z500 en JJA sur l'hémisphère Nord et l'hémisphère Sud. On constate comme dans le cas de l'hiver boréal une réduction des biais de Z500 avec l'introduction de la dynamique stochastique aléatoire. Toutefois cette réduction est moins marquée, et la distribution des biais de Z500 est plus zonale initialement. Il n'y a donc pas lieu de s'attendre à un impact important de cette réduction de biais sur la circulation atmosphérique à l'échelle intra-saisonnière.

4.3.2.4 Comparaison avec la méthode de nudging d'anomalie

Afin de comparer les deux méthodes de perturbations stochastiques étudiées plus haut, les scores des ensembles de prévisions avec perturbations issues du nudging itéré sont calculés sur la période commune de 1989–2010 pour la saison DJF. La figure 4.8 montre l'évolution mois par mois de l'erreur quadratique moyenne (après avoir retiré le biais moyen en validation

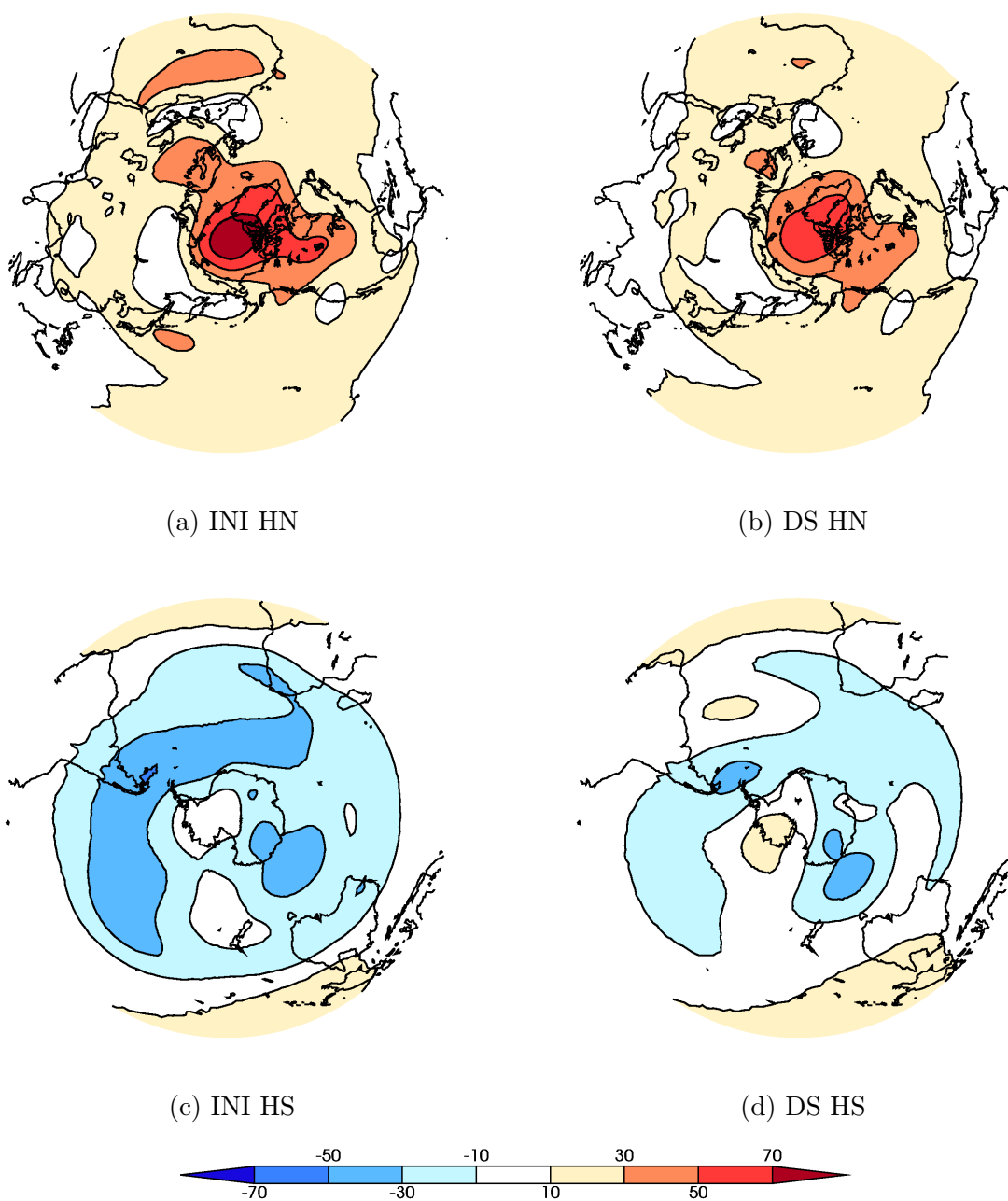


FIGURE 4.7 – Biais (en m) sur l'hémisphère Nord (en haut) et l'hémisphère Sud (en bas) du Z500 de la saison JJA des ensembles CM5 INI et DS RAND.

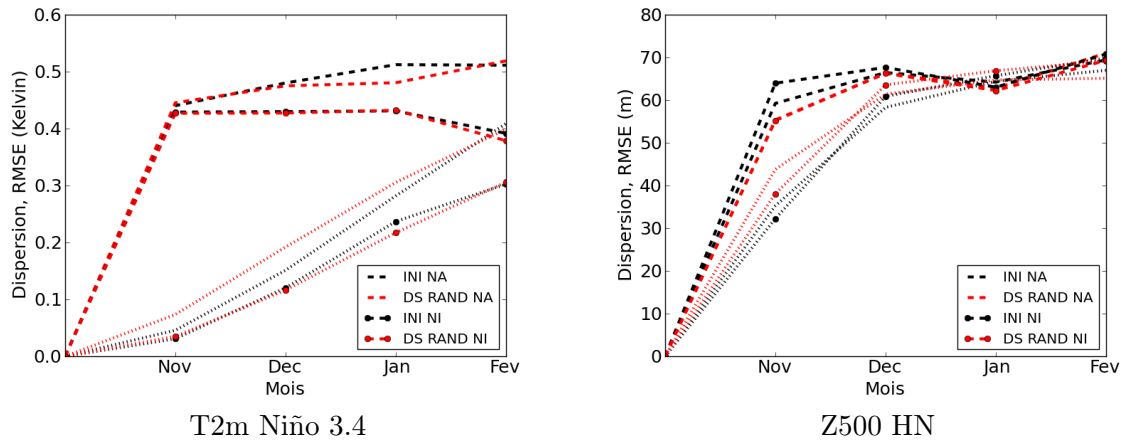


FIGURE 4.8 – Evolution mois par mois de la dispersion d’ensemble (pointillés) et de l’erreur quadratique moyenne (RMSE, tirets épais) au cours des prévisions de NDJF 1989–2010 de T2m sur la région Niño 3.4 (à gauche) et de Z500 sur l’hémisphère Nord (à droite). Les ensembles étudiés sont INI (en noir) et DS RAND (en rouge) pour les méthodes basées sur le nudging d’anomalie (NA) ou nudging simple itéré (NI).

croisée) et de la dispersion des ensembles, avec et sans dynamique stochastique, et utilisant la méthode de nudging d’anomalie (NA) ou de nudging itéré (NI). Sur la région Niño 3.4, les prévisions basées sur le nudging itéré ont une erreur quadratique moindre, mais aussi moins de dispersion que celles basées sur le nudging d’anomalie, ce dès le premier mois. Contrairement aux ensembles NA, le RMSE n’augmente pas. Il diminue même au quatrième mois de la prévision. Aucun impact de la dynamique stochastique sur la dispersion d’ensemble de la température de surface n’est visible sur cette région, alors qu’il est assez net pendant les trois premiers mois de prévision pour les ensembles NA. De manière générale, le rapport erreur/dispersion des prévisions de température de surface sur la région Niño 3.4 est amélioré avec la méthode basée sur le nudging d’anomalie, tandis que celle basée sur le nudging simple itéré est sans effet.

Pour le Z500 sur l’hémisphère Nord, les contrastes sont plus nets pour le premier mois de prévision que le pour le reste de la saison entre les deux méthodes NA et NI, et l’introduction ou non de dynamique stochastique. Les deux courbes de RMSE avec dynamique stochastique aléatoire sont quasiment superposées et en dessous de celles sans dynamique stochastique. Pour le premier mois, c’est la méthode de dynamique stochastique basée sur le nudging d’anomalie qui donne le meilleur rapport erreur-dispersion, tandis que par la suite ce sont plutôt les deux ensembles de la méthode basée sur le nudging itéré.

La figure 4.9 montre l’évolution mois par mois des scores de mACC des quatre ensembles pour la T2m sur la région Niño 3.4. Les prévisions des expériences NI sont meilleures que NA, mais au cours des trois premiers mois des prévisions, la dynamique stochastique permet à DS RAND NA d’avoir une décroissance du mACC plus faible que les autres ensembles. Par contre, cet effet disparaît au cours du mois de février et le score de mACC chute à 0.85.

On peut comparer les scores de corrélation des anomalies en recalculant des fourchettes de mACC pour 9 membres parmi 15 sur la période commune 1989–2010 pour les ensembles basés sur la méthode de nudging itéré. Les conclusions des comparaisons entre les ensembles (CM5) INI NI et DS RAND NI sur la période complète 1979–2010 tiennent toujours sur la

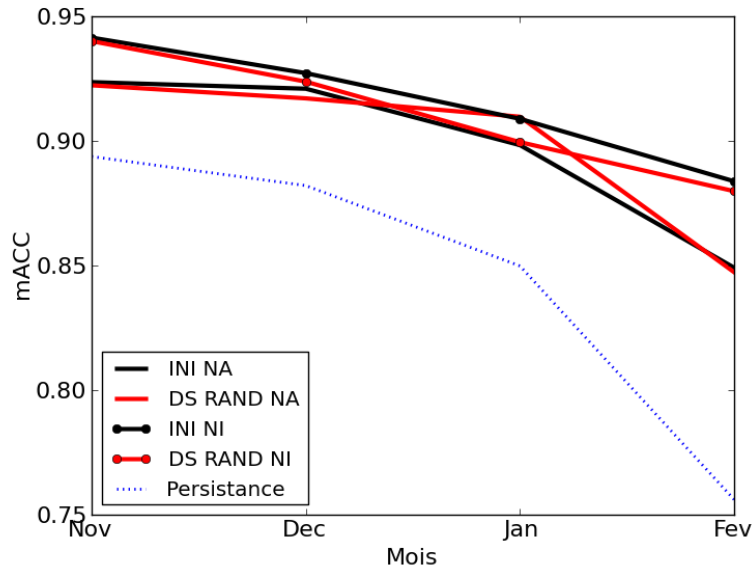


FIGURE 4.9 – Evolution mois par mois de la mACC des prévisions de T2m sur la région Niño 3.4 pour NDJF 1989–2010 avec les ensembles INI (en noir) et DS RAND (en rouge) pour les méthodes basées sur le nudging d’anomalie (NA) ou nudging simple itéré (NI). Les scores de la persistance des anomalies au premier novembre sont également montrés (pointillés bleus).

période restreinte à 1989–2010. Par rapport aux ensembles INI NA et DS RAND NA, les scores de mACC sont significativement meilleurs pour les deux ensembles NI pour la T2m sur l’hémisphère Nord, les tropiques et la région Niño 3.4, ainsi que pour les précipitations sur les tropiques. Les mACC du Z500 sur l’hémisphère Nord sont significativement améliorés par rapport à DS RAND NA pour l’ensemble DS RAND NI, tandis que INI NA est significativement meilleur que INI NI.

Un autre score pouvant préciser les différences entre les jeux de prévisions sur l’hémisphère Nord en hiver est la corrélation des séries d’indices climatiques. On s’intéresse donc ici aux corrélations entre les prévisions et les réanalyses ERA-Interim des indices de NAO et NAM pour les saisons DJF 1989–2010. Les résultats obtenus sont listés dans le tableau 4.14. Deux conclusions découlent de ces corrélations et de leurs fourchettes : les scores de NAM et de NAO sont très variables selon les membres des ensembles, avec des fourchettes de scores parfois très larges, mais l’ensemble DS RAND NI est significativement meilleur que les trois autres pour la prévision du NAM et du NAO sur la période d’étude.

L’ensemble de ces scores et comparaisons illustrent la complémentarité des deux méthodes. En somme, la méthode basée sur le nudging d’anomalie semble améliorer les prévisions sur le Pacifique tropical mais dégrade les scores sur l’hémisphère Nord en hiver, tandis que la méthode basée sur le nudging itéré permet d’améliorer ces scores sur ces régions mais dégrade légèrement les scores sur la région Niño 3.4. Pour des raisons de rapidité de calcul des prévisions saisonnières, les méthodes développées ici reposent sur l’ajout de champs spectraux (donc globaux) en guise de perturbations. On peut se demander dans quelle mesure des perturbations ciblées (sur les régions tropicales, ou sur les latitudes tempérées) pourraient permettre d’améliorer les prévisions, sans oublier toutefois le fait que ces régions sont reliées entre elles par d’importantes téléconnexions.

Indice	INI NA			DS RAND NA		
	5%	95%	ACC	5%	95%	ACC
NAO	0.02	0.31	0.16	-0.10	0.35	0.12
NAM	-0.39	0.31	-0.02	-0.16	0.33	0.11
Indice	INI NI			DS RAND NI		
	5%	95%	ACC	5%	95%	ACC
NAO	0.05	0.37	0.20	0.29	0.48	0.39
NAM	-0.07	0.12	0.03	0.23	0.51	0.37

TABLEAU 4.14 – Scores de corrélation des indices de NAO et NAM pour DJF 1989–2010 des ensembles INI et DS RAND conçus à partir du nudging d’anomalie (NA, en haut) ou du nudging simple itéré (NI). Le tableau donne la moyenne des ACC de tirages de 9 membres parmi 15, et les fourchettes de 5% à 95% ainsi obtenues.

4.3.3 Impact des corrections moyennes

Les résultats des paragraphes précédents ont montré que l’inclusion de la dynamique stochastique avec des tirages aléatoires des perturbations définies par un nudging itéré améliore de manière significative les prévisions du Z500 sur l’hémisphère Nord en hiver. Nous avons également vu, en décomposant la moyenne quadratique du terme de perturbation, que la part correspondant aux perturbations moyennes est faible par rapport à la variance haute fréquence. Dans quelle mesure les améliorations évoquées sont-elles dues à la moyenne des perturbations plutôt qu’à l’introduction de bruit dans les membres de la prévision ? Pour répondre à cette question, une rétro-prévision MMENS est calculée. Chaque membre de cet ensemble est perturbé à $t = 0$ comme CM5 INI, et à chaque pas de temps, un terme correspondant à la moyenne pour le mois courant M des $\{\delta\mathbf{X}\}_M$ des autres années de la période d’étude est ajouté à la dynamique du modèle ARPEGE-Climat.

4.3.3.1 Saison DJF

Les scores de mACC de MMENS pour DJF sont montrés avec ceux des autres ensembles associés au nudging itéré dans le tableau 4.10 (page 128). Le mACC sur la région Niño 3.4 est encore dégradé par rapport à l’expérience DS RAND, de sorte que les fourchettes de scores sont disjointes. Sur les tropiques, les scores de T2m sont significativement moins bons que DS RAND et équivalents à CM5 INI, tandis que pour les précipitations, une nette dégradation des scores est observée par rapport à toutes les autres méthodes testées. Par contre, sur l’hémisphère Nord, MMENS est significativement meilleur que CM5 INI pour Z500 et équivalent pour la T2m.

Pour ce qui est des scores probabilistes, la décomposition en fiabilité et résolution et le RPSS de MMENS sont donnés dans le tableau 4.15. Les scores de RPSS, à comparer à ceux du tableau 4.11, confirment les améliorations et dégradations notées pour le mACC par rapport aux ensembles CM5 INI et DS RAND. Si la différence du RPSS sur la région Niño 3.4 entre CM5 INI et DS RAND était due à une perte de fiabilité, le RPSS de MMENS est nettement plus faible du fait d’une forte dégradation du terme de résolution. Sur l’hémisphère Nord, le RPSS est encore amélioré par rapport à DS RAND, et cette amélioration est due à une meilleure résolution sans dégrader la fiabilité par rapport à CM5 INI.

Ensemble		MMENS			DS MENS			DS SEQ5		
Paramètre		$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS	$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS	$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS
T2m	Niño 3.4	0.299	0.892	0.593	0.293	0.894	0.601	0.294	0.912	0.618
Précip.	Trop.	0.468	0.512	0.044	0.451	0.501	0.051	0.451	0.504	0.052
Z500	HN	0.399	0.434	0.035	0.391	0.413	0.022	0.392	0.419	0.027
T2m	HN	0.462	0.490	0.028	0.460	0.474	0.014	0.458	0.486	0.028

TABLEAU 4.15 – Idem tableau 4.11, pour les expériences MMENS, DS MENS et DS SEQ5 et la saison DJF 1979–2010.

Ensemble		MMENS			DS MENS			DS SEQ5		
Paramètre		$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS	$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS	$\frac{Rel}{Unc}$	$\frac{Res}{Unc}$	RPSS
T2m	Niño 3.4	0.411	0.765	0.354	0.424	0.769	0.345	0.402	0.770	0.368
Précip.	Trop.	0.502	0.530	0.028	0.506	0.528	0.022	0.496	0.525	0.029
Z500	HN	0.391	0.386	-0.006	0.379	0.385	0.006	0.382	0.379	-0.003
T2m	HN	0.409	0.408	-0.001	0.398	0.404	0.005	0.400	0.401	0.002

TABLEAU 4.16 – Idem tableau 4.15, pour la saison JJA 1979–2010.

Dans le cas des prévisions d’hiver sur les latitudes tempérées, c’est donc en grande partie la correction moyenne des erreurs de tendance du modèle apportée par les perturbations stochastiques qui est à l’origine de l’amélioration des scores du Z500. Par contre, sur les tropiques, les améliorations des scores constatées avec DS RAND sont dues à l’aspect stochastique des perturbations, tandis que la dégradation des scores sur la région Niño 3.4 est due à cette correction systématique, et atténuée par l’aspect stochastique des perturbations dans le cas de DS RAND.

4.3.3.2 Saison JJA

Pour la saison JJA, les scores de mACC sont donnés dans le tableau 4.12 et les scores de RPSS dans le tableau 4.16. Les corrélations d’anomalies sont quasiment identiques à celles obtenues avec DS RAND, si ce n’est dans le cas de Z500 sur l’hémisphère Sud où le mACC ne peut plus être distingué de CM5 INI, et est significativement meilleur que DS RAND. Dans ce cas on voit donc que l’ajout de perturbations de haute fréquence dégrade les perturbations. A l’inverse, sur l’hémisphère Nord en été, le mACC est plus dégradé par l’ajout du terme moyen mensuel que par les perturbations aléatoires.

Les scores de RPSS sont encore dégradés par rapport à DS RAND (voir le tableau 4.13). Pour la T2m sur la région Niño 3.4 et Z500 sur l’hémisphère Nord, ces dégradations sont essentiellement le fait d’une moins bonne fiabilité.

4.3.4 Quelle fréquence des perturbations ?

Dans le paragraphe précédent, nous avons vu que la perturbation du modèle avec un terme moyen constant impactait les prévisions d’hiver boréal de manière similaire à la dynamique stochastique aléatoire sur les moyennes latitudes, mais que les prévisions sur les tropiques étaient dégradées. Plusieurs expériences ont été menées afin de mieux déterminer le rôle des

différentes échelles de temps des perturbations dans l'amélioration ou la dégradation des prévisions d'ensemble.

4.3.4.1 Corrections mensuelles aléatoires

Dans un premier temps, le rôle de la variabilité interannuelle des corrections a été étudié en regardant l'impact de corrections mensuelles aléatoires sur les prévisions. Pour ce faire, les moyennes mensuelles des corrections d'erreur de tendance initiale pour les quatre membres de la prévision nudgée ont été calculées chaque année de la période de rétro-prévision. Ceci donne pour chaque mois M et année y une population de 124 corrections possibles correspondant aux 31 autres années de la période 1979–2010. Chaque mois, un tirage sans remise d'une perturbation par membre est effectué, et cette perturbation est ajoutée à la dynamique du modèle tout au long du mois.

Les ensembles définis de cette manière sont calculés pour les saisons DJF et JJA et notés DS MENS.

4.3.4.2 Corrections séquentielles

La décomposition des termes quadratiques moyens a montré l'importance de la variabilité intra-annuelle des corrections d'erreur de tendance initiale. En tirant aléatoirement une nouvelle perturbation chaque jour, il est possible que certains effets des corrections soient atténués, par compensation de perturbations de signe opposé. Une piste d'amélioration de la dynamique stochastique consisterait donc à choisir des perturbations cohérentes dans le temps pendant une période donnée, en tirant non plus une nouvelle perturbation par membre chaque jour, mais en tirant tous les n jours une séquence de n perturbations quotidiennes consécutives, provenant d'un membre de la prévision nudgée et d'une autre année de la période de rétro-prévision. L'échelle synoptique de 5 jours a été retenue ici. Les rétro-prévisions d'hiver et d'été sont notées DS SEQ5.

4.3.4.3 Résultats obtenus

Les mACC des prévisions déterministes de ces deux expériences sont comparés à ceux des autres ensembles basés sur le nudging itéré dans les tableaux 4.10 (page 128) et 4.12 (page 130) pour DJF et JJA, respectivement. Les résultats des prévisions probabilistes en termes de scores RPSS et leur décomposition en fiabilité et résolution sont montrés dans les tableaux 4.15 pour DJF et 4.16 pour JJA.

◦ Saison DJF :

En DJF, la prévision DS MENS obtient des scores très similaires à MMENS pour la T2m sur les tropiques et la région Niño 3.4, tandis que le mACC de Z500 et de la T2m sur l'hémisphère Nord sont significativement moins bons. Le mACC de Z500 sur l'hémisphère Sud est inchangé, tandis que les précipitations sur les tropiques ont un score similaire à CM5 INI et donc significativement meilleur que MMENS.

Sur les différentes régions étudiées en DJF, l'ensemble DS SEQ5 donne des scores de mACC parmi les deux meilleurs des 5 ensembles comparés dans le tableau 4.10, hormis pour la T2m

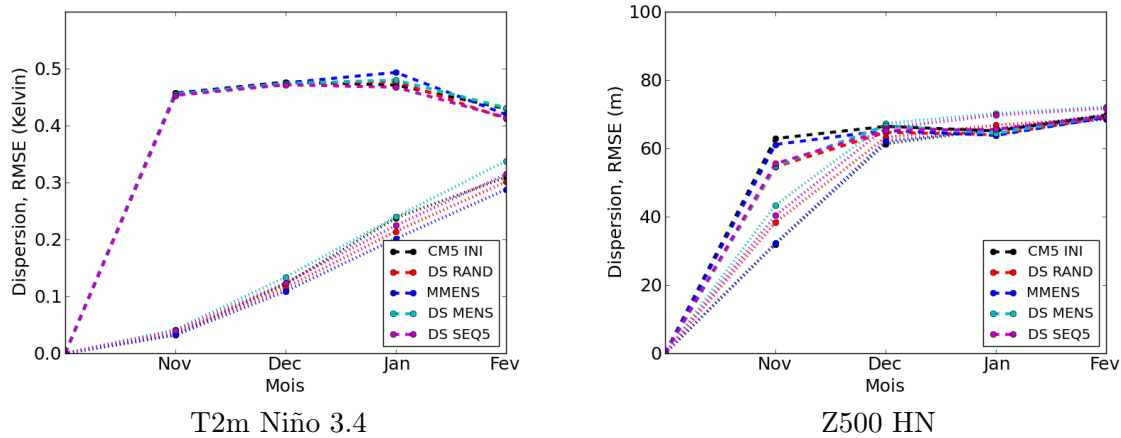


FIGURE 4.10 – Idem figure 4.8, pour NDJF 1979–2010 et les ensembles basés sur le nudging itéré pour la définition des perturbations.

sur les tropiques (mais les écarts pour ce score entre les différents ensembles ne sont pas significatifs). Pour la saison DJF, les tirages de perturbations séquentielles donnent donc les meilleurs résultats pour la dynamique stochastique, et améliorent les prévisions déterministes par rapport à CM5 INI pour les précipitations tropicales et le Z500 sur l’hémisphère Nord. Le seul bémol est la réduction significative du score de mACC de T2m sur la région Niño 3.4, constatée avec tous les ensembles incluant la dynamique stochastique.

Sur la figure 4.10, les courbes de dispersion des différents ensembles pour le Z500 sur l’hémisphère Nord en DJF et T2m sur la région Niño 3.4 montrent que la dispersion d’ensemble des deux variables est la plus élevée pour l’ensemble DS MENS. Sur la région Niño 3.4, cette dispersion est à peine plus élevée que celle de CM5 INI, tandis que celle des autres ensembles avec perturbations est la plupart du temps inférieure. Les deux figures illustrent très clairement que la méthode DS SEQ5 (en magenta) est intermédiaire entre une méthode purement aléatoire (DS RAND, en rouge) et une méthode à l’échelle mensuelle (DS MENS, en cyan). Les résultats pour Z500 montrent que l’amélioration du rapport erreur/dispersion n’est pas nécessairement garant d’une amélioration des prévisions : DS MENS et DS SEQ5 ont des comportements très similaires chaque mois de la période NDJF, mais des scores de prévision de Z500 sur la région significativement différents.

Le tableau 4.17 liste les scores de corrélation des anomalies des indices de NAO et NAM des cinq ensembles basés sur le nudging itéré étudiés dans ce chapitre. Les ACC des indices NAO et NAM sont significativement améliorés avec les méthodes DS RAND et MMENS par rapport aux perturbations initiales, tandis que la méthode DS MENS a peu d’impact sur les scores obtenus comme c’est le cas pour le mACC de Z500 sur l’hémisphère Nord. Le cas de DS SEQ5 est assez singulier, dans la mesure où le score moyen d’ACC est amélioré par rapport à CM5 INI et DS MENS, mais la fourchette des scores est tellement large qu’on ne peut pas conclure à une amélioration significative.

Les scores de RPSS pour DJF sont montrés dans le tableau 4.15 et peuvent être comparés à ceux de CM5 INI et DS RAND donnés dans le tableau 4.11. Le RPSS de l’ensemble DS MENS est meilleur que MMENS sur les tropiques, mais dégradé sur l’hémisphère Nord. DS MENS a systématiquement une meilleure fiabilité que MMENS, mais la résolution de l’ensemble est

Ensemble	NAO			NAM		
	5%	95 %	ACC	5%	95%	ACC
CM5 INI	0.00	0.25	0.12	-0.01	0.22	0.10
DS RAND	0.23	0.41	0.32	0.23	0.43	0.34
MMENS	0.13	0.42	0.28	0.17	0.39	0.28
DS MENS	0.00	0.22	0.11	-0.02	0.26	0.12
DS SEQ5	0.01	0.40	0.20	-0.06	0.40	0.18

TABLEAU 4.17 – Fourchettes à 5% et 95% des ACC de 9 membres parmi 15 des indices NAO et NAM pour DJF 1979–2010 des différents ensembles basés sur le nudging itéré.

dégradée sauf dans le cas de la T2m sur la région Niño 3.4. DS MENS et DS SEQ5 ont des scores de fiabilité très similaires, parmi les meilleurs des 5 ensembles étudiés dans cette partie. La résolution est par contre nettement améliorée avec DS SEQ5 par rapport à DS MENS. Cet ensemble donne les scores de RPSS parmi les deux meilleurs de tous les ensembles pour les régions étudiées, hormis Z500 sur l'hémisphère Nord où il se classe troisième. Le RPSS de DS SEQ5 pour la T2m sur la région Niño 3.4 est même légèrement meilleur que celui de CM5 INI.

◦ Saison JJA :

En JJA les conclusions sont légèrement différentes. Le mACC de la T2m sur la région Niño 3.4 est encore dégradé avec DS MENS par rapport à MMENS, faisant de DS MENS le moins bon ensemble sur la région. Les scores sur les tropiques sont quasiment identiques entre les deux méthodes, tandis qu'on retrouve sur l'hémisphère Sud en JJA la dégradation du mACC de Z500 observée sur l'hémisphère Nord en DJF avec DS MENS. L'amélioration du mACC de Z500 sur l'hémisphère Nord avec DS MENS est toutefois significative, et le score obtenu est semblable à CM5 INI. Le mACC de la T2m sur cette région est également significativement amélioré par rapport à MMENS et le score moyen dépasse celui de CM5 INI.

L'ensemble DS SEQ5 n'obtient pas des résultats aussi nets vis-à-vis des autres méthodes qu'en DJF. Sur les régions tropicales, les mACC de DS SEQ5 sont systématiquement parmi les deux meilleurs ensembles. C'est le cas également pour la T2m sur l'hémisphère Nord où l'ensemble est au même niveau que CM5 INI. Par contre, pour Z500 sur les deux régions tempérées HN et HS, il y a peu de différences entre DS RAND et DS SEQ5, qui est donc significativement moins bon que CM5 INI.

En comparant les scores de RPSS obtenus avec ces deux méthodes aux autres ensembles issus de la méthode de nudging itéré (tableaux 4.13 et 4.16), il apparaît que DS MENS est le moins bon des ensembles sur les régions tropicales, mais a des scores similaires à CM5 INI sur les extra-tropiques et surclasse donc les autres méthodes de perturbations. Par rapport à MMENS, la fiabilité des prévisions de Z500 et T2m sur l'hémisphère Nord est améliorée, tandis que la résolution est légèrement moins bonne. Comme c'est le cas pour les scores de mACC, DS SEQ5 est le meilleur des modèles avec dynamique stochastique pour T2m sur la région Niño 3.4, et très proche de CM5 INI pour la T2m sur l'hémisphère Nord. Hormis le cas de T2m sur la région Niño 3.4, les scores du tableau 4.16 montrent que la fiabilité des prévisions de DS SEQ5 est en général meilleure ou équivalente à celle des autres ensembles avec dynamique stochastique, mais que la résolution est légèrement moins bonne. Les scores de RPSS sont donc du même ordre que les autres ensembles avec dynamique stochastique, et

Durée (en jours)						
Régime	ERA-I	CM5 INI	DS RAND	MMENS	DS MENS	DS SEQ5
DA	5.89	6.03	6.06	6.54	6.53	6.18
Zonal	9.60	7.12	8.32	8.36	8.27	7.75
Blocage	7.38	6.60	7.53	7.00	7.49	7.11
NAO ⁻	9.15	8.31	9.48	8.96	9.73	9.19
Fréquence (en %)						
Régime	ERA-I	CM5 INI	DS RAND	MMENS	DS MENS	DS SEQ5
DA	16	20	18	19	18	18
Zonal	33	28	30	31	29	29
Blocage	25	22	23	22	23	23
NAO ⁻	19	15	18	18	19	19
ACC						
Régime	ERA-I	CM5 INI	DS RAND	MMENS	DS MENS	DS SEQ5
DA	-	0.05	-0.20	0.21	0.04	-0.08
Zonal	-	0.24	0.25	0.19	0.02	-0.19
Blocage	-	-0.01	0.19	-0.22	-0.23	-0.21
NAO ⁻	-	-0.06	-0.14	0.08	-0.10	0.20
BSS	-	0.09	0.03	0.02	-0.07	0.03

TABEAU 4.18 – De haut en bas : durée moyenne (en jours) et fréquence (en % de jours) des quatre régimes de temps sur la région Atlantique Nord - Europe pour la période DJF 1979–2010 dans les réanalyses ERA-Interim (ERA-I) et les cinq ensembles de prévisions saisonnières basés sur le nudging itéré ; scores de corrélation des anomalies saisonnières de fréquence de ces régimes et Brier Skill Score sur cette même période des ensembles basés sur le nudging itéré vis-à-vis des données ERA-Interim.

aucun impact particulièrement marqué du tirage de perturbations séquentielles n'est visible en JJA sur les résultats en termes de prévisions probabilistes.

4.3.4.4 Impact sur la variabilité intra-saisonnière

On peut évaluer l'impact de la dynamique stochastique sur la variabilité intra-saisonnière du climat à travers le prisme des régimes de temps sur la région Atlantique-Nord Europe. Les différentes caractéristiques des régimes de temps (durée et fréquence) ainsi que les scores d'ACC et de BSS des ensembles basés sur le nudging itéré sont montrés dans le tableau 4.18.

La durée et la fréquence des régimes sont assez similaires entre les données des réanalyses ERA-Interim et les différents ensembles de rétro-prévisions. Toutefois, l'ensemble CM5 INI sous-estime les durées et fréquences des régimes Zonal, Blocage et NAO⁻ par rapport à ERA-Interim, et l'introduction de la dynamique stochastique permet de corriger en partie ces manquements. Ceci n'induit pas une amélioration des scores des rétro-prévisions, puisque le BSS et la moyenne des scores de corrélation des anomalies sont dégradés avec l'introduction de la dynamique stochastique dans le modèle. Il y a toutefois une très grande variabilité des scores d'ACC entre les différents ensembles. C'est également le cas entre les différents membres des ensembles de rétro-prévision, et il est impossible de conclure quant à la supériorité d'un ensemble par rapport à un autre. Il est intéressant ici de remarquer que les scores obtenus paraissent nettement en deçà des scores montrés dans le paragraphe 4.2.3.3 pour les ensembles basés sur le nudging d'anomalie. En calculant les scores des ensembles basés sur le nudging

itéré sur la période 1989–2010, on obtient des scores sensiblement équivalents (mais légèrement inférieurs) à ceux montrés pour les ensembles INI NA et DS NA. Cet exemple illustre la difficulté de l'évaluation des rétro-prévisions saisonnières mais aussi l'importance de l'utilisation de périodes de rétro-prévision les plus longues possibles pour la robustesse des résultats.

En termes de climat moyen sur la période de réanalyse, les résultats obtenus montrent peu d'impact de la fréquence des perturbations sur la fréquence des différents régimes de temps. Par contre, les tirages séquentiels avec une fréquence de 5 jours (ensemble DS SEQ5) donnent des régimes généralement plus courts que ceux observés avec une méthode aléatoire avec une fréquence de tirage de 6 heures (DS RAND) ou mensuelle (DS MENS). Dans le cas des régimes Blocage et Zonal, ceci correspond à une amélioration du climat moyen du modèle, tandis qu'au contraire, pour le régime NAO⁻, l'ensemble DS SEQ5 se rapproche des données de référence.

4.4 Synthèse

Ce chapitre a permis d'introduire la méthode de « dynamique stochastique » développée au cours de cette thèse, et d'évaluer l'impact de plusieurs variantes de cette méthode sur les prévisions saisonnières déterministes et probabilistes. Les choix d'implémentation faits dans le cadre de ces travaux ont défini une technique de perturbation de la dynamique du modèle d'atmosphère ARPEGE-Climat hybride entre des perturbations stochastiques additives et des corrections d'erreurs de tendance initiale des variables pronostiques du modèle.

Deux moyens de déterminer la population des perturbations additives du modèle $\delta\mathbf{X}$ ont été explorés dans ce chapitre. Tous deux reposent sur la technique de nudging. La méthode de nudging d'anomalie (NA) permet théoriquement de garder le modèle couplé CNRM-CM5.1 à l'équilibre, et d'obtenir des perturbations centrées. Celle basée sur un nudging simple itéré (NI) permet de tenir compte de l'évolution des erreurs de tendance initiale du modèle en fonction du délai de la prévision saisonnière, et de corriger les écarts du modèle aux données de réanalyse ERA-Interim. Ces deux méthodes de perturbations aléatoires en cours d'intégration du modèle (DS RAND) ont été comparées à une méthode de perturbations initiales, consistant à ajouter au temps $t = 0$ uniquement une perturbation aléatoire $\delta\mathbf{X}$ à chaque membre.

Les résultats obtenus peuvent être résumés comme suit : pour la saison DJF, la dynamique stochastique basée sur le nudging d'anomalie (NA) permet d'améliorer la prévision de la température de surface sur la région Niño 3.4, et est équivalente aux perturbations initiales pour les précipitations et la T2m sur les tropiques, tandis que les scores déterministes et probabilistes des prévisions sur l'hémisphère Nord sont dégradés. On retrouve des résultats similaires en JJA, si ce n'est que les prévisions déterministes sur les tropiques sont significativement améliorées avec la dynamique stochastique. La dynamique stochastique basée sur le nudging itéré (NI) permet au contraire d'améliorer les prévisions sur les régions tempérées en DJF. Les résultats sur les tropiques sont très similaires entre les méthodes de dynamique stochastique aléatoire et de perturbations initiales, tandis que les prévisions de l'ENSO sont dégradées par la dynamique stochastique.

Le mode d'obtention des perturbations est donc un facteur déterminant pour l'impact de la dynamique stochastique sur les prévisions. A ce titre, la décomposition des termes de perturbations quadratiques moyennes a montré l'existence d'un terme moyen non nul non seulement dans le cas du nudging itéré (où ce terme est attendu) mais également dans le cas du nudging d'anomalie (où initialement, nous nous attendions à avoir des perturbations

centrées). Pour le nudging d'anomalie, l'existence de ce terme démontre que la dérive initiale du modèle nudgé n'est pas la même qu'en cours de prévision. Contrairement au nudging itéré, où ce terme moyen va corriger en partie des biais du modèle par rapport à la climatologie des données de référence, le terme moyen en mode nudging d'anomalie ne correspond pas à un écart du modèle à la réalité en cours de prévision. Ceci peut expliquer les dégradations des prévisions sur les latitudes tempérées dans ce cas.

Une expérience consistant à corriger à chaque pas de temps la dynamique du modèle par la moyenne mensuelle sur la période de réanalyse des corrections (en validation croisée) illustre l'importance du terme de correction moyenne pour l'amélioration des prévisions sur les latitudes tempérées de l'hémisphère Nord en hiver (et leur dégradation en été). À l'inverse, les résultats de cette expérience montrent l'importance du terme aléatoire de haute fréquence sur les régions tropicales. Deux jeux de prévisions perturbées avec une fréquence différente permettent d'affiner les résultats précédents, en tirant des perturbations moyennes à l'échelle mensuelle (DS MENS), et des perturbations séquentielles à l'échelle synoptique de 5 jours (DS SEQ5). L'ensemble DS MENS est le plus dispersif, mais imposer la même correction tout au long d'un mois donne des résultats mitigés, avec une nette dégradation des prévisions sur les régions tempérées.

La solution de perturber le modèle avec une séquence cohérente de perturbations (provenant de 5 jours de rappels consécutifs du modèle vers les données de réanalyse lors de l'intégration nudgée du modèle) donne dans l'ensemble les meilleurs résultats parmi les différentes versions de la dynamique stochastique présentées dans ce chapitre. Cette solution paraît également satisfaisante sur le plan théorique, dans la mesure où elle intègre une quatrième dimension cohérente aux perturbations ajoutées aux variables pronostiques du modèle. Forts de ce constat, nous pouvons pousser plus loin cette idée de cohérence des perturbations, et étudier si une classification des perturbations du modèle en fonction de l'état de celui-ci au cours de l'intégration nudgée ne permettrait pas d'améliorer les prévisions saisonnières. Les séquences de cinq jours de perturbations sont pour l'instant tirées de manière purement aléatoire parmi les autres années de la période de rétro-prévision, mais cette échelle suggère par exemple l'utilisation de régimes de temps pour définir dans quelle sous-population tirer des perturbations du modèle. Ce conditionnement des perturbations fait l'objet du chapitre suivant.

Chapter Summary

This chapter introduces the stochastic dynamics technique developed during this thesis, and quantifies the impact of several versions of this method on deterministic and probabilistic seasonal prediction skill. This technique is designed as a perturbation method that acts on the atmospheric component ARPEGE-Climat dynamics. It can be seen as a hybrid between an additive stochastic perturbation method and an initial tendency error correction method.

Two means of defining the population of additive model perturbations $\delta\mathbf{X}$ were tested. Both rely on model nudging. With the anomaly nudging technique (NA) the coupled model CNRM-CM5.1 can theoretically stay close to its balanced state, and perturbations are then centered. An iterative nudging technique (NI) takes into account the evolution of model initial tendency errors with lead-time in a seasonal forecast, and tends to correct the model discrepancies with respect to the ERA-Interim reanalysis data. The two in-run random perturbation techniques (DS RAND) were compared to initial perturbation ensembles.

Results can be summarized as follows : during DJF, the NA stochastic dynamics technique improves surface temperature forecasts over the Niño 3.4 region, and shows skill similar to the initial perturbations ensemble for precipitation and 2-meter temperature over the tropics, whereas deterministic and probabilistic forecast skill is worsened over the Northern Hemisphere mid-latitudes. Similar results are found in JJA, save for deterministic forecasts over the tropics which are significantly improved with stochastic dynamics. On the other hand, the NI stochastic dynamics technique improves DJF predictions over mid-latitudes and exhibits skill quite similar to the initial perturbations ensemble over the tropics, save for ENSO predictions.

The way perturbations are designed is therefore a very important point to take into account. Splitting the mean square perturbation terms into a time-average squared perturbation term, interannual, and intra-annual variance terms shows that the time average term is non-zero, not only in the case of NI but also for NA (for which perturbations are expected to be centered). This means that the initial drift of the model when nudged is different from initial errors in seasonal prediction mode. Unlike the NI method where the average term can correct some model biases with respect to reference data, the average correction term in the case of NA accounts for nothing real. This could explain why seasonal predictions are worse over mid-latitudes with NA stochastic dynamics.

Different sets of re-forecasts using alternative stochastic dynamics techniques based on NI investigate sources of improvement. A monthly mean in-run correction technique shows that over the Northern Hemisphere mid-latitudes, the average correction corresponding to the calendar month (in cross-validation mode) has a greater impact on results than the stochasticity of corrections. Opposite conclusions can be drawn from results over the tropics. Drawing mean perturbations at a monthly time scale increases ensemble dispersion but hinders ensemble skill over mid-latitudes, whereas drawing perturbations corresponding to a sequence of five days of corrections gives (on average) the best results for the different methods tested in this chapter.

Using a sequence of perturbations seems also satisfactory on a theoretical basis, since this method adds a fourth coherent dimension to the stochastic additive perturbations. This conclusion calls for further investigation. The idea of classifying perturbations, depending on the state of the model during the run, is addressed in the next chapter.

VERS UN CONDITIONNEMENT DES PERTURBATIONS

Le chapitre précédent a permis de montrer qu’une méthode de perturbations stochastiques de la dynamique du modèle ARPEGE-Climat, appelée « dynamique stochastique », peut améliorer certains aspects des prévisions saisonnières avec le modèle couplé CNRM-CM5.1.

Dans ce chapitre, on cherche à affiner cette technique, en se posant plusieurs questions fondamentales :

- Les perturbations du modèle telles que définies par la technique de dynamique stochastique contiennent-elles une information utile à l’échelle interannuelle ?
- Que peut-on alors espérer améliorer en extrayant la bonne information de ces perturbations ?
- Comment peut-on conditionner les $\{\delta\mathbf{X}\}$ de manière à tirer des perturbations dans un sous-ensemble adapté à l’état du modèle en cours de prévision ?
- Le conditionnement permet-il d’améliorer les prévisions ?

Afin de répondre à ces questions, ce chapitre est divisé en plusieurs parties. En premier lieu, des rétro-prévisions saisonnières ont été implémentées avec perturbations « optimales », c’est à dire en partant de la connaissance des données à prévoir pour le mois en cours.

Ensuite, plusieurs méthodes de conditionnement des perturbations ont été définies, implémentées et testées en prévision saisonnière dans le cadre de cette thèse. Elles sont basées pour l’une sur les TSO sur la région Niño 3.4, et pour les autres sur les champs globaux de fonction de courant du modèle ARPEGE-Climat. Les résultats sont présentés dans les parties 5.2 et 5.3 de ce chapitre.

5.1 Quelles améliorations possibles ? Perturbations optimales à l'échelle mensuelle

Dans cette partie nous verrons que les perturbations d'une année donnée comportent une information susceptible d'améliorer sensiblement les prévisions saisonnières, ceci indépendamment de la séquence des perturbations à l'échelle intra-mensuelle.

La dynamique stochastique basée sur le nudging itéré et l'impact des tirages de perturbations optimales sur les scores de la saison DJF ont fait l'objet d'un article publié dans la revue *Geophysical Research Letters* en 2012, par lequel nous commençons cette partie. L'étude est ensuite étendue à une comparaison des deux méthodes (nudging itéré et nudging d'anomalie) sur la période commune de rétro-prévisions, et pour les deux saisons DJF et JJA.

Les résultats obtenus dans cette partie motivent les études de conditionnement des perturbations montrées dans la suite du chapitre.

5.1.1 Résumé de l'article publié dans *Geophysical Research Letters*

Dans cet article est décrite la méthode de dynamique stochastique basée sur le nudging itéré étudiée dans la partie 4.3 du chapitre précédent. Les résultats obtenus pour la saison d'hiver sur les latitudes tempérées sont montrés pour deux ensembles similaires à CM5 INI et DS RAND¹.

L'originalité de l'article repose sur l'intégration de jeux de prévisions avec perturbations stochastiques « optimales », consistant à tirer des perturbations correspondant aux erreurs de tendance initiale estimées pour le mois en cours de prévision (*ie* janvier 1990 pour janvier 1990). Par contre les perturbations ne correspondent pas au jour en question, ce qui permet de créer un ensemble de prévisions avec une dispersion raisonnable. L'idée est d'estimer de manière empirique la borne supérieure des améliorations qu'on peut espérer d'une méthode de conditionnement optimale des perturbations à l'échelle mensuelle. Bien entendu, cette méthode ne peut être utilisée en temps réel. Les scores obtenus sont un indicateur des variables et régions pour lesquelles des progrès sont possibles en améliorant la technique de dynamique stochastique, même s'il est illusoire d'attendre d'une technique de conditionnement des perturbations des améliorations aussi conséquentes.

Les résultats de l'article montrent que la corrélation de l'indice NAO de l'ensemble avec perturbations optimales DS OPT atteint 0.71 et que les scores de corrélation des anomalies de Z500, température et précipitations sont significativement améliorés sur l'hémisphère Nord et les tropiques. Le mode de tirage des perturbations de l'ensemble DS OPT ne peut évidemment pas être appliqué pour des prévisions opérationnelles. Par contre, ces résultats suggèrent qu'un conditionnement des perturbations à l'échelle mensuelle pourrait améliorer les prévisions, principalement sur les régions tempérées.

1. On notera que ces ensembles, ainsi que DS OPT, ont été calculés à partir d'états initiaux océaniques nudgés, et non NEMOVAR comme ceux étudiés dans la partie 4.3, mais que ce choix a peu d'impact sur les résultats obtenus pour la saison DJF.

5.1.2 Article : A stochastic method for improving seasonal predictions

By Lauriane BATTÉ and Michel DÉQUÉ *Météo France, CNRM-GAME, 42 avenue G. Coriolis, FR-31057 Toulouse Cedex, France*

(Received 17 February 2012; accepted 3 April 2012; published 4 May 2012)

Citation : L. Batté and M. Déqué (2012). A stochastic method for improving seasonal predictions. *Geophys. Res. Lett.*, 39, L09707. doi : 10.1029/2012GL051406.

5.1.2.1 Abstract

Ensemble seasonal forecasts during boreal winter suffer from insufficient spread and systematic errors. In this study we suggest a new stochastic dynamics method to address both issues at a time. Our technique relies on random additive corrections of initial tendency error estimates of the atmospheric component of the CNRM-CM5.1 global climate model, using ERA-Interim as a reference over a 1979-2010 hindcast period. The random method improves deterministic scores for 500-hPa geopotential height forecasts over the Northern Hemisphere extratropics (NH Z500), and increases the ensemble spread. An optimal method consisting in drawing the error corrections within the current month of the hindcast period illustrates the high potential of future improvements, with NH Z500 anomaly correlation reaching 0.65 and North Atlantic Oscillation index correlation 0.71 with ERA-Interim. These substantial improvements using current year corrections pave the way for future forecasting methods using classification criteria on the correction population.

Published in Geophysical Research Letters. Copyright 2012 American Geophysical Union. Further reproduction or electronic distribution is not permitted.

5.1.2.2 Introduction

Seasonal prediction using coupled general circulation models (GCMs) has been an active field of research over the last two decades. International research efforts such as the European Commission-funded DEMETER (Palmer et al., 2004) and ENSEMBLES (Doblas-Reyes et al., 2009; Weisheimer et al., 2009) projects as well as the APEC Climate Center-sponsored CLIPAS project (Wang et al., 2009) illustrated the potential of state-of-the-art numerical climate models in forecasting temperature and geopotential, and to a lesser extent precipitation, at a seasonal timescale. Predictability is generally higher over the Tropics, but models show positive skill with respect to climatology over some midlatitudinal regions. Most model ensembles suffer from systematic errors and lack of spread. Multi-model techniques pooling together predictions from several models address both issues : some systematic errors are cancelled out provided that individual model errors are different, and reliability is improved (Hagedorn et al., 2005). However, the success of a multi-model ensemble technique relies mainly on the quality of the individual models used. In addition, if a model has insufficient spread and a large prediction error over a given region, it will lead the multi-model towards a wrong prediction.

In recent years a variety of stochastic perturbation methods has been implemented in atmospheric models to account for model error, both for short-term ensemble predictions and monthly-to-seasonal forecasts using these models as the atmospheric component of an

earth-system model. Buizza et al. (1999) introduced random perturbations of model physical tendencies into the ECMWF ensemble prediction system. An additional scheme called Stochastic Kinetic Energy Backscatter (SKEB) algorithm is used by ECMWF to scatter kinetic energy dissipated by the model at the sub-grid scale back to larger scales (Shutts, 2005), and Berner et al. (2008) highlights the reduction of systematic error and improvements of most deterministic and probabilistic skill scores over different regions at a seasonal time scale due to this algorithm. SKEB is used alongside a perturbed parameters scheme described in Bowler et al. (2008) in the Met Office’s GloSea4 seasonal forecast model (Arribas et al., 2011). Similar stochastic physics schemes are also used for medium-range forecasts in the Canadian ensemble prediction system (Charron et al., 2010).

In the present study, an alternative stochastic perturbation technique is applied to the CNRM-CM5.1 GCM (Voldoire et al., 2012) for seasonal forecasting. Predictions are stochastically corrected by adding randomly drawn initial tendency residuals to the temperature, specific humidity and vorticity fields in the prognostic equations of the ARPEGE-Climat v5.2 atmospheric model component. The initial tendency residuals are estimated using a nudging technique as described in Kaas et al. (1999) and Guldberg et al. (2005). Several past studies such as Yang and Anderson (2000), Barreiro and Chang (2004) and Guldberg et al. (2005) have suggested that correcting systematic errors in atmospheric or coupled ocean-atmosphere GCMs reduce model bias with some impact on seasonal prediction skill. However, Guldberg et al. (2005) found that systematic error correction in a previous version of ARPEGE-Climat showed no improvement over the Tropics and the Northern Hemisphere. The originality of the method presented here relies on the stochasticity of the error corrections. A more detailed description of the stochastic dynamics technique is given in section 2, and results are shown in section 3. They illustrate the significant gain in seasonal forecasting skill during Northern Hemisphere winter. An upper limit for possible future improvements using this method is also shown.

5.1.2.3 Stochastic Dynamics Method

The stochastic dynamics method implemented in the ARPEGE-Climat v5.2 atmospheric model for seasonal forecasts is an additive stochastic perturbation of three prognostic ARPEGE variables \mathbf{X} : temperature, specific humidity and vorticity, following equation 5.1. $\mathbf{M}(\mathbf{X}(t), t)$ represents the evolution of variable \mathbf{X} due to the initial ARPEGE-Climat model formulation, and $\delta\mathbf{X}_t$ is the stochastic perturbation.

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \mathbf{M}(\mathbf{X}(t), t) + \delta\mathbf{X}_t \quad (5.1)$$

Our method derives from Guldberg et al. (2005) and consists in using the nudging technique to estimate initial tendency errors of ARPEGE-Climat v5.2 and then perturbing a seasonal forecast with random initial tendency error corrections drawn within these estimates. The stochastic dynamics method follows three steps. The first step is to run the CNRM-CM5.1 model during 32 years (1979-2010), nudging it towards the ECMWF ERA-Interim reanalysis data (Dee et al., 2011). ERA-Interim data is re-interpolated on the ARPEGE-Climat reduced gaussian grid. Prognostic variables temperature, specific humidity and vorticity are relaxed towards the ERA-Interim fields with relaxation times of a day for temperature and specific humidity and 6 hours for vorticity. This run provides initial conditions on November 1st 1979

to 2010 (for boreal winter forecasts) for each component of CNRM-CM5.1.

In a second step, a four-member ensemble is implemented for each November-December-January-February season (NDJF) of the 1979-2010 period. This second run is relaxed more weakly towards ERA-Interim and started with initial conditions from the first run, thus reducing spin-up effects due to differences between ERA-Interim and model climatology. Relaxation times are selected close to one month for temperature and specific humidity, and ten days for vorticity. A vertical profile for relaxation coefficients is introduced in the five lowest levels of the model so as to tune relaxation down to zero and avoid inconsistencies at the surface. Differences between ERA-Interim fields and each member for the three relaxed variables are stored daily. The opposite of these fields, thus corresponding to model corrections towards ERA-Interim, make up the $\{\delta\mathbf{X}\}$ population from which the perturbations are drawn in forecast mode.

The third step consists in the actual retrospective forecast, started with initial conditions each November 1st from the first run and with perturbations drawn from the $\{\delta\mathbf{X}\}$ population designed in the second step of the method. In this study perturbations were drawn within the corresponding calendar month, meaning that $\{\delta\mathbf{X}\}$ was in fact separated in four bins for NDJF coherent with the forecast lead-time. A different $\delta\mathbf{X}$ was drawn for each ensemble member every six hours of the forecast. Perturbations for temperature, specific humidity and vorticity are drawn together, and correspond to an error correction for a given day of the second step re-forecast. This ensures that perturbations are coherent between the three corrected fields, and avoids partially cancelling out the effects of one correction with that of another field.

5.1.2.4 Experiments and Results

Three sets of seasonal re-forecasts of December to February (DJF) 1979-80 to 2010-11 were run with 15 ensemble members :

1. The reference seasonal forecast ensemble (REF) was perturbed with random $\delta\mathbf{X}$ drawn from the initial tendency error correction population only at the initial time step.
2. A random stochastic dynamics ensemble (SD_RANDOM) was perturbed with $\delta\mathbf{X}_t$ at each time step.
3. An optimal stochastic dynamics ensemble (SD_OPT) was perturbed with $\delta\mathbf{X}_t$ at each time step drawn in the same month and year as the actual forecast.

The SD_OPT experiment cannot be implemented for operational forecasts, since initial tendency errors can only be estimated for a set of hindcasts. Perturbations are consistent with the errors the model makes in a given month. Therefore, results for SD_OPT determine the upper limit for scores using this stochastic perturbation technique, provided that corrections are relevant to the model initial tendency errors at a given time.

The impact of the stochastic dynamics method on DJF 500 hPa geopotential height (Z500) bias over the Northern Hemisphere is shown in figure 5.1. The negative bias over the polar region is reduced in SD_RANDOM, and Z500 bias gradients over the northern Pacific and northern Atlantic are less pronounced. SD_OPT biases are very similar to SD_RANDOM (not shown). Figure 5.2 shows anomaly correlation coefficients (ACC) for DJF Z500 over the Northern Hemisphere extra-tropics (30 to 75 degrees North) for each forecast ensemble. The random

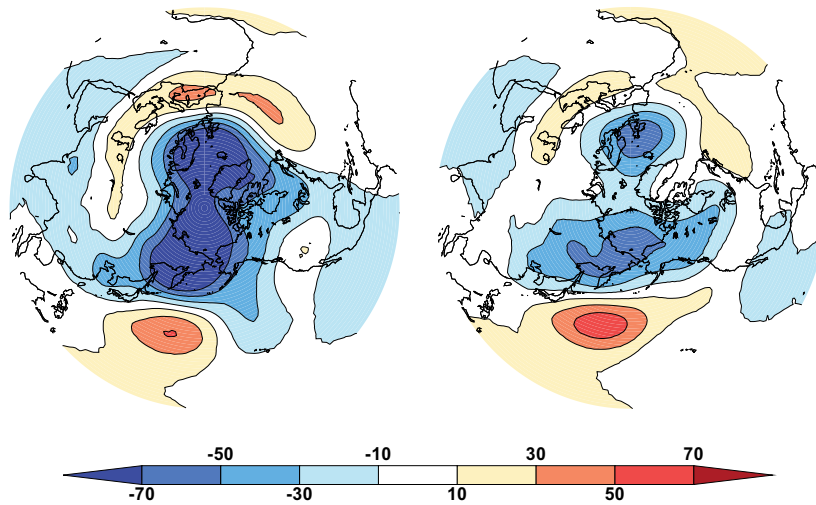


FIGURE 5.1 – DJF NH Z500 mean bias (in meters) for ensembles REF (left) and SD_RANDOM (right).

TABLEAU 5.1 – Mean ACC values for REF, SD_RANDOM and SD_OPT. Statistical significance of differences between the SD ensembles and REF are tested using a binomial test for season ACC scores. Bold scores are significantly better than REF at a 95% level.

Region	Variable	REF	SD_RANDOM	SD_OPT
NH ^a	Z500	0.25	0.37	0.65
Tropics ^b	Precipitation	0.45	0.45	0.52
Tropics	T2m	0.47	0.47	0.51
Niño 3.4 ^c	T2m	0.83	0.81	0.82

^a 30°N-75°N

^b 23°N-23°S

^c 170°W-120°W and 5°N-5°S

stochastic dynamics method improves anomaly correlation for 22 out of 32 seasons. The associated binomial test shows that this improvement is statistically significant ($p = 0.025$). While the REF ensemble yields correlation values lower than 0.2 for 15 seasons, correlation remains lower than this threshold for only 8 seasons with the SD_RANDOM ensemble. SD_OPT anomaly correlation scores reach over 0.6 for 19 seasons and are lower than 0.4 for only 4 seasons. This suggests that an appropriate set of perturbations in a given season could lead to significant improvements in forecasting skill.

Mean ACC values for different variables and regions were calculated for the three ensembles and are listed in table 5.1. Mean ACC is considerably improved with stochastic dynamics for Z500 over the Northern Hemisphere extra-tropics, in coherence with results shown earlier. Results over the Tropics for 2-meter temperature (T2m) and precipitation and the Niño 3.4 region for T2m exhibit no significant impact of the stochastic dynamics method on mean ACC scores for SD_RANDOM, whereas SD_OPT improves precipitation and T2m scores over the Tropics.

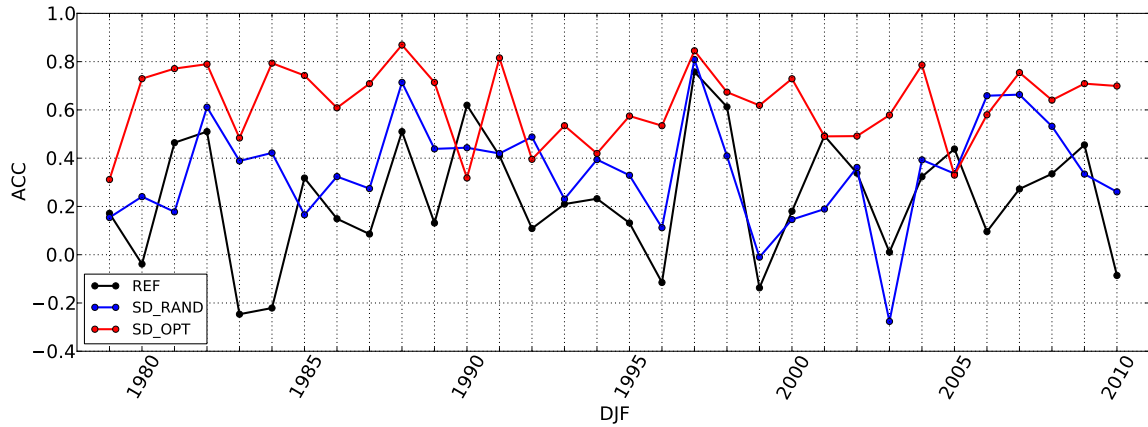


FIGURE 5.2 – DJF NH Z500 anomaly correlation coefficient for ensembles REF, SD_RANDOM and SD_OPT.

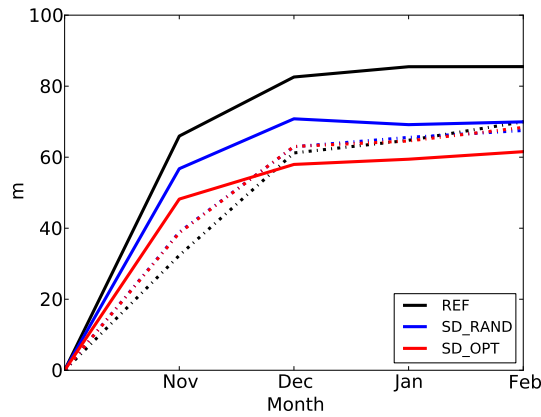


FIGURE 5.3 – Evolution of monthly root mean square error (full lines) and ensemble spread (dashed lines) for NH Z500 with forecasts REF, SD_RANDOM and SD_OPT.

Improvement over the Northern Hemisphere extra-tropics is also found when looking at monthly root mean square error (RMSE) of the forecasts over the 1979-2010 time period. Figure 5.3 illustrates the improvement of the spread-to-skill ratio of the forecast ensemble for NH Z500. While RMSE is reduced by over 15 meters in months 3 and 4 of the forecast, the SD_RANDOM ensemble also has a higher spread during the first two months, and similar spread in the following two months. The stochastic dynamics method therefore improves model error and dispersion, as intended. SD_OPT has the same spread as SD_RANDOM, with an ensemble spread larger than the RMSE after a 2-month lead.

Skill was further assessed over the Euro-Atlantic region by investigating model performance in forecasting the North Atlantic Oscillation (NAO). Following a method similar to Doblas-Reyes et al. (2003), the NAO is defined as the leading empirical orthogonal function (EOF) of December to February monthly Z500 ERA-Interim data from 1979 to 2010 over the region 20°N-80°N and 80°W-40°E. Model NAO indexes are calculated by projecting monthly grid point anomalies for each member onto this EOF. Forecasts and ERA-Interim verification series are standardized in cross-validation mode. The introduction of stochastic dynamics has little

TABLEAU 5.2 – Reliability, resolution, RPS and RPSS values for ERA-Interim climatology, REF, SD_RAND and SD_OPT for NH T2m (land grid points only) and Z500. Bold RPS values indicate scores significantly better than REF at a 95% level using a binomial test for season RPS scores.

Ensemble	Rel	Res	RPS	RPSS
NH T2m (over land)				
Climatology	0.	0.	0.222	-
REF	0.095	0.099	0.218	0.019
SD_RAND	0.094	0.100	0.217	0.026
SD_OPT	0.094	0.112	0.204	0.080
NH Z500				
Climatology	0.	0.	0.222	-
REF	0.090	0.095	0.217	0.022
SD_RAND	0.088	0.097	0.213	0.042
SD_OPT	0.091	0.120	0.193	0.131

impact on the ensemble spread of the forecasts at a seasonal time scale. The SD_RAND ensemble has slightly higher skill than REF in forecasting the NAO index, with a correlation of 0.36 versus 0.32 between the ensemble mean index and the reference ERA-Interim index. The SD_OPT ensemble exhibits significant improvement with a correlation of 0.71 with ERA-Interim.

Probabilistic skill was evaluated with a ranked probability score (RPS) for tercile prediction defined following Toth et al. (2003) as the average of Brier Scores for a given variable remaining below the climatological terciles. The RPS ranges between 0 (perfect forecast) and 1 and consists in a sum over the 32 seasons of quadratic distances in probabilistic space between forecasts and observations (worth 0 or 1 whether the event occurs or not a given season). Reliability, resolution (Murphy, 1973) and RPS scores are calculated as in Batté and Déqué (2011) for each grid point over land and averaged over the region of interest. Results for T2m terciles over NH land grid points and NH Z500 are shown in table 5.2. A ranked probability skill score is defined as $RPSS = 1 - RPS/RPS_c$ where RPS_c is the climatology RPS. Similar scores are found for ensembles REF and SD_RAND, which outperform climatological forecasts over the region, yielding positive RPSS values. The improvement in scores noted for SD_OPT is mainly due to an increase in resolution, which evaluates the ability of the model to separate events that have different probabilities of occurrence.

5.1.2.5 Conclusion and Discussion

This study presents an original technique for stochastic perturbations combining the assets of random perturbation and systematic error correction in coupled models used for seasonal forecasts. Re-forecasts of DJF 1979-2010 using this method with the CNRM-CM5.1 GCM show enhanced performance over the Northern Hemisphere for 500hPa geopotential height, with similar skill over the Tropics. RMSE and anomaly correlation coefficients for Z500 show that random stochastic perturbations as designed in our study can enhance scores and improve the model spread-to-skill ratio. These improvements are triggered by a reduced seasonal bias consistent with previous studies that corrected average errors, and an enhanced ensemble spread consistent with other stochastic techniques.

Results with an ensemble using optimal corrections drawn from the current forecast month suggest room for improvement in seasonal forecasting skill, provided that corrections are drawn from a population that is representative of the common initial tendency errors of the current season. Correlation coefficients for the NAO index with the optimal ensemble reach 0.7 and therefore illustrate the potential of such a technique, as long as an appropriate classification of the correction population is found. Further work should therefore focus on exploring classification criteria for the perturbation population based on the state of the ocean or the atmosphere, using analogues to classify perturbations according to tropical sea surface temperature or weather regimes as in D’Andrea and Vautard (2000). It is worth mentioning that although RMSE was further reduced with optimal perturbations, ensemble spread remained very close to the random perturbation ensemble. A concise study of probabilistic skill showed that ranked probability score improvements with the optimal ensemble relied mainly on increased resolution. Lack of improvement in reliability could be corrected by multi-model forecasting. Given the current impact of our method on model spread, other stochastic perturbations with a longer time scale could be included in the model. Future experiments will study the impact of the perturbation frequencies and drawing several successive chronological corrections on model spread and skill.

ACKNOWLEDGMENTS

ERA-Interim data used in this study were provided by ECMWF. We are grateful to two anonymous reviewers who helped us improve this manuscript.

5.1.3 Cas des perturbations issues du nudging d’anomalie

Les perturbations issues du nudging itéré contiennent un terme de correction des erreurs du modèle par rapport aux données nous servant de référence, à savoir les réanalyses ERA-Interim.

Dans le cas des perturbations issues du nudging d’anomalie, les termes $\delta\mathbf{X}$ sont (en théorie) centrés par construction, et nous avons vu dans le chapitre 4 que la part de la variabilité interannuelle était faible dans la décomposition du terme quadratique moyen des perturbations. On peut donc se demander quelles seraient les performances d’un ensemble DS OPT dans ce cas. Un ensemble d’expériences noté OPT NA a été implémenté et comparé à l’ensemble OPT NI pour les saisons d’été et d’hiver sur la période de réanalyse commune 1989–2010.

5.1.3.1 Résultats de la saison DJF et comparaison avec le nudging itéré

L’ensemble OPT NA pour DJF comporte 15 membres, pour lesquels les perturbations correspondent au même mois de l’intégration nudgée que celui en cours de rétro-prévision. De la même manière que DS RAND, une nouvelle perturbation est tirée toutes les 6 heures.

La figure 5.4 montre les scores de mACC et leurs fourchettes de 5% à 95% pour 9 membres parmi 15 sur différentes régions et pour les paramètres T2m, précipitations et Z500. Les scores des ensembles INI, DS RAND et OPT issus du nudging d’anomalie (NA) et du nudging itéré (NI) sont calculés sur la période commune des rétro-prévisions DJF 1989–2010.

La comparaison entre les méthodes NA et NI ayant déjà été abordée dans le chapitre précédent, nous décrirons surtout ici les apports relatifs des perturbations stochastiques dites

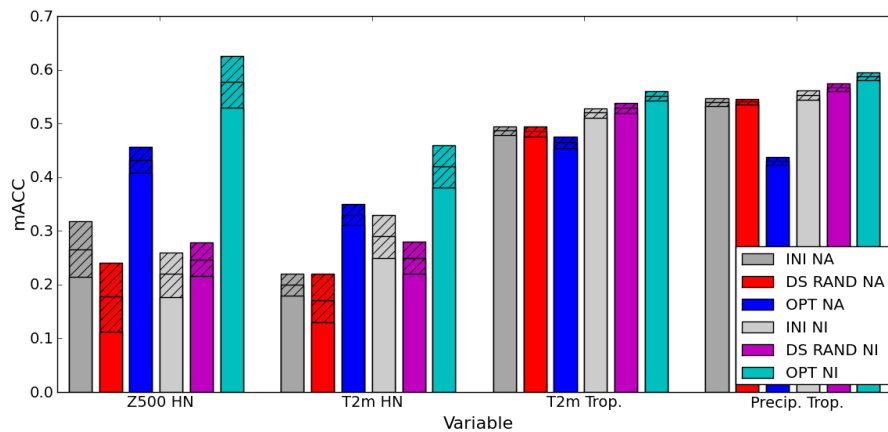


FIGURE 5.4 – Scores de mACC et fourchettes de 5% à 95% de tirages de 9 membres parmi 15 des ensembles INI, DS RAND et OPT basés sur des perturbations issues du nudging d’anomalie (NA) et du nudging itéré (NI), pour DJF 1989–2010 et différentes régions et paramètres étudiés.

« optimales », par rapport à des perturbations initiales ou une dynamique stochastique purement aléatoire pour chacune des méthodes. Pour les perturbations NA, on note que les perturbations optimales permettent d’améliorer de manière considérable les scores de mACC sur l’hémisphère Nord extra-tropical. La corrélation des anomalies de Z500 est augmentée de plus de 140% par rapport à DS RAND et de plus de 60% par rapport à INI. Par contre, sur les régions tropicales, les perturbations optimales dégradent les scores obtenus. Ce résultat surprenant est examiné plus en détail par la suite.

Les résultats montrés pour la méthode NI sur la période 1989–2010 sont cohérents avec ceux publiés dans *Geophysical Research Letters*, à savoir que les corrélations des anomalies sont améliorées significativement pour l’ensemble des paramètres sur les régions étudiées, mais que le gain est beaucoup plus marqué sur les régions tempérées que sur les tropiques.

Une question importante qui survient lors de l’évaluation de la méthode optimale est celle de l’évolution du rapport erreur/dispersion du modèle. On peut s’attendre à ce que la dispersion d’ensemble soit inférieure à celle obtenue avec la méthode DS RAND, puisque les perturbations proviennent d’une population plus petite, et en théorie plus cohérente. Toutefois, des résultats précédents ont déjà montré que certains effets de la dynamique stochastique sont non-linéaires. La figure 5.5 montre les résultats obtenus pour Z500 sur l’hémisphère Nord. Il apparaît sur cette figure que les perturbations optimales issues du nudging itéré permettent de très nettement réduire l’erreur quadratique moyenne du modèle sur la région. C’est également le cas pour le nudging d’anomalie, mais de manière moins marquée. En termes de dispersion, il y a très peu de différences entre les méthodes DS RAND et OPT correspondantes au cours des deux premiers mois de prévision, tandis que la dispersion des prévisions OPT est légèrement plus faible aux mois 3 et 4.

Les résultats de la partie 4.3.2.4 du chapitre précédent ont montré que la dynamique stochastique pouvait améliorer les prévisions d’ensemble des indices NAO et NAM, du moins avec la méthode NI. Le potentiel des perturbations optimales pour la prévision de ces indices est mis en évidence dans le tableau 5.3. Les améliorations des scores de corrélation des indices NAO et NAM sont significatives pour les deux méthodes OPT NA et OPT NI par rapport

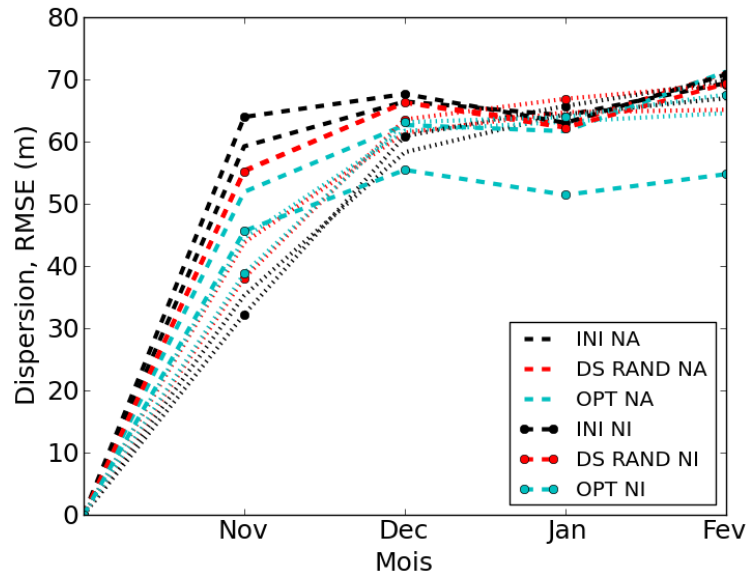


FIGURE 5.5 – Evolution mois par mois de la dispersion d'ensemble (pointillés) et de l'erreur quadratique moyenne (RMSE, tirets épais) au cours des prévisions de NDJF 1989–2010 de Z500 sur l'hémisphère Nord. Les ensembles étudiés sont INI (en noir), DS RAND (en rouge) et OPT (en cyan) pour les méthodes basées sur le nudging d'anomalie (NA) ou nudging simple itéré (NI).

	OPT NA			OPT NI		
Indice	5%	95%	ACC	5%	95%	ACC
NAO	0.73	0.80	0.77	0.64	0.82	0.74
NAM	0.64	0.85	0.76	0.40	0.67	0.54
	INI NA			INI NI		
Indice	5%	95%	ACC	5%	95%	ACC
NAO	0.02	0.31	0.16	0.05	0.37	0.20
NAM	-0.39	0.31	-0.02	-0.07	0.12	0.03
	DS RAND NA			DS RAND NI		
Indice	5%	95%	ACC	5%	95%	ACC
NAO	-0.10	0.35	0.12	0.29	0.48	0.39
NAM	-0.16	0.33	0.11	0.23	0.51	0.37

TABLEAU 5.3 – Scores de corrélation des indices de NAO et NAM pour DJF 1989–2010 des ensembles OPT. On rappelle les scores des ensembles INI et DS RAND du tableau 4.14 du chapitre précédent.

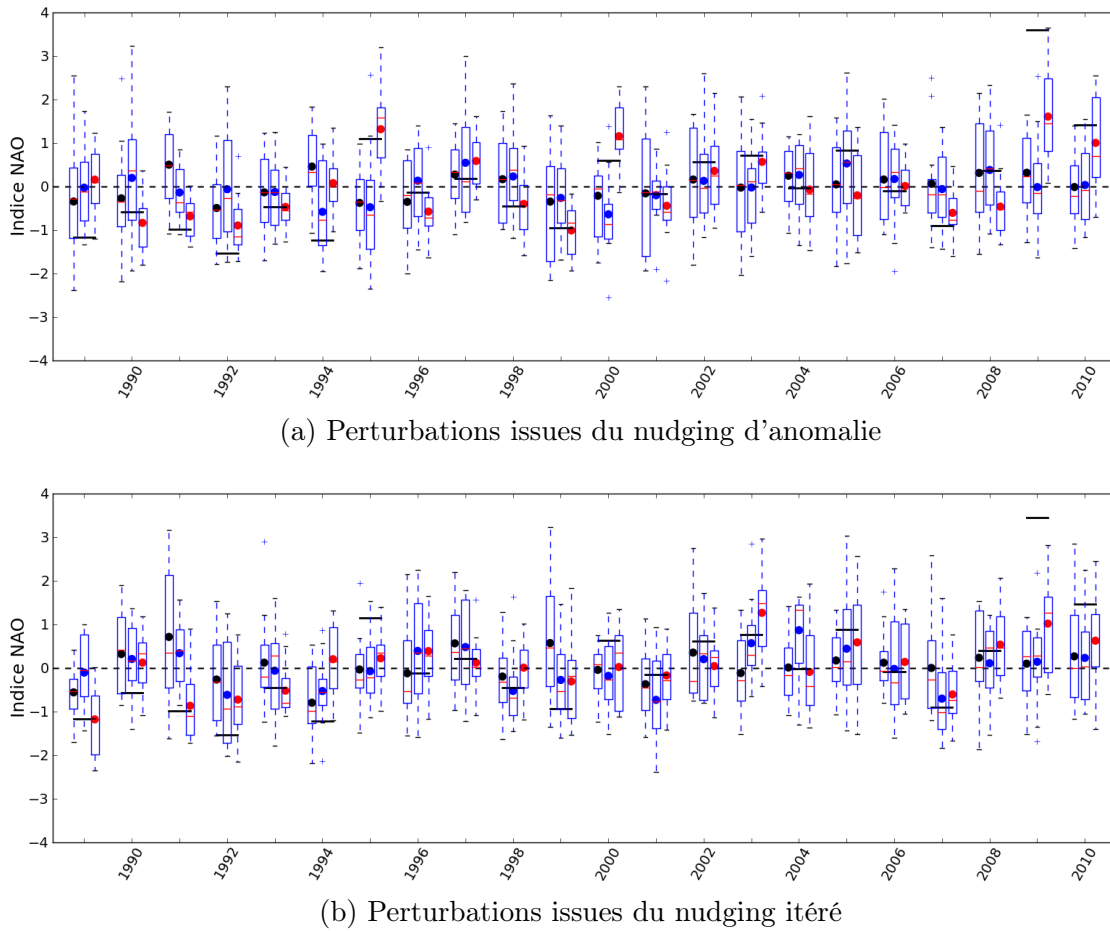


FIGURE 5.6 – Boîtes à moustaches et moyenne d'ensemble des indices NAO des prévisions de DJF de 1989 à 2010 avec INI (noir), DS RAND (bleu) et OPT (rouge) pour les méthodes NA (a) et NI (b) ; les indices NAO observés sur la période sont montrés par les barres noires épaisses.

aux ensembles INI et DS RAND respectifs. Les figures 5.6 (a) et (b) montrent les indices de NAO moyens et la dispersion d'ensemble des expériences INI, DS RAND et OPT pour les méthodes NA et NI, respectivement, sur la période commune DJF 1989–2010. La variabilité interannuelle des indices NAO prévus par OPT NA est légèrement plus élevée que celle de OPT NI. Notons également que l'événement particulièrement marqué de l'hiver 2009–2010 est mieux représenté par OPT NA, pour lequel tous les membres prévoient un indice NAO positif, ce qui n'arrive que pour cette année sur la période de rétro-prévision. La figure 5.6 montre également qu'il y a peu d'impact de l'introduction des perturbations stochastiques sur la dispersion d'ensemble des prévisions d'indice NAO. La dispersion est légèrement réduite en moyenne dans le cas des perturbations optimales, mais ceci est loin d'être systématique pour les différentes années montrées ici.

5.1.3.2 Résultats de la saison JJA

De la même manière que les prévisions de DJF, on compare les ensembles INI, DS RAND et OPT des méthodes NA et NI sur la période commune JJA 1989–2010. Les résultats en

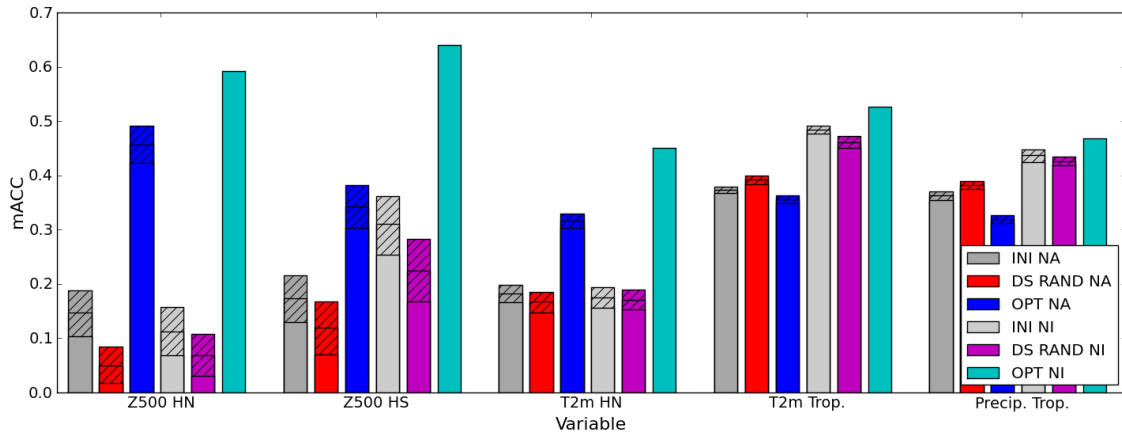


FIGURE 5.7 – Idem figure 5.4, pour JJA 1989–2010. L’absence de fourchettes pour OPT NI provient du fait que c’est un ensemble de 9 membres seulement.

termes de mACC sont montrés dans la figure 5.7².

Les conclusions sont très similaires à celles obtenues pour DJF. Si les perturbations optimales NI permettent d’améliorer significativement les scores de mACC sur les régions étudiées pour les paramètres T2m, Z500 et précipitations, l’amélioration est beaucoup plus marquée sur les régions tempérées que sur les tropiques. Avec les perturbations issues du nudging d’anomalie (NA), les mACC sont encore légèrement (mais significativement) dégradés sur les tropiques.

5.1.3.3 Pourquoi une dégradation des scores sur les tropiques ?

On peut s’étonner dans le cas du nudging d’anomalie de l’asymétrie entre tropiques et latitudes tempérées dans les impacts des perturbations optimales sur les scores de mACC, d’autant plus que l’introduction de perturbations stochastiques aléatoires pour les prévisions de la saison JJA a plutôt tendance à dégrader les scores de mACC sur les latitudes tempérées sans impacter les tropiques. Revenons à la décomposition de la moyenne des carrés des perturbations calculée au chapitre précédent (figure 4.1, page 115). Pour la plupart des variables, la part de la variance interannuelle dans cette décomposition est plus importante sur les tropiques en hiver que sur les régions tempérées. Par contre, les corrections moyennes de fonction de courant représentent une plus grande part de cette décomposition sur les régions tempérées. Toutefois, on ne peut rien conclure directement de cette décomposition tant les deux régions sont liées, et les effets observés sur les tropiques peuvent être liés aux corrections sur les latitudes tempérées et inversement.

Afin de mieux comprendre la nature des améliorations et dégradations de mACC, l’évolution temporelle de la corrélation des anomalies spatiales est étudiée pour les différentes régions et variables. La figure 5.8 montre les résultats obtenus pour Z500 sur l’hémisphère Nord et pour la température de surface et les précipitations sur les tropiques. L’amélioration des scores d’ACC pour Z500 HN est observée quasiment chaque année par rapport à DS RAND (19 an-

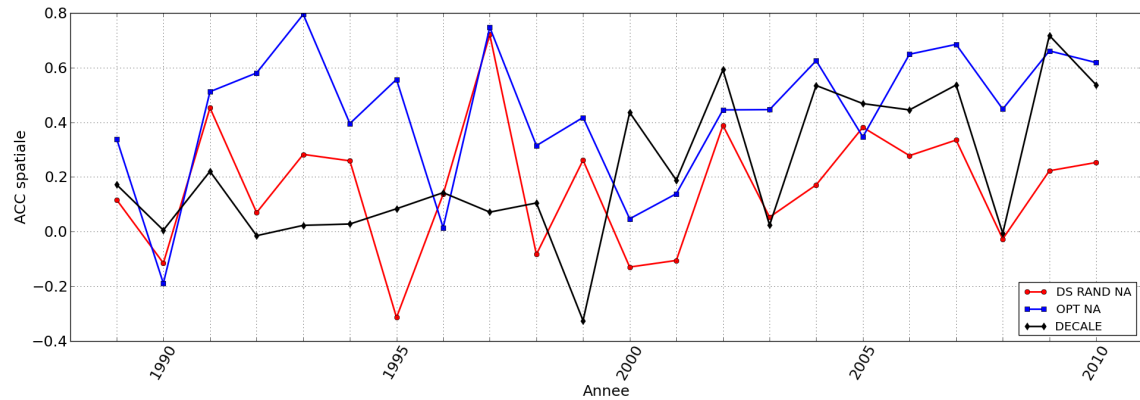
2. Au vu des résultats obtenus avec perturbations optimales pour la saison d’hiver, il n’a pas été jugé opportun d’implémenter 15 membres pour l’ensemble OPT NI

nées sur les 22 de la période). Les biais de Z500 des ensembles DS RAND NA et OPT NA pour les trois années dégradées par les perturbations optimales, ne mettent pas en évidence de différence marquée avec les biais de trois années pour lesquelles l'ACC spatiale est la plus améliorée. L'effet des perturbations optimales n'est donc pas uniquement un effet de correction des biais du modèle une année donnée.

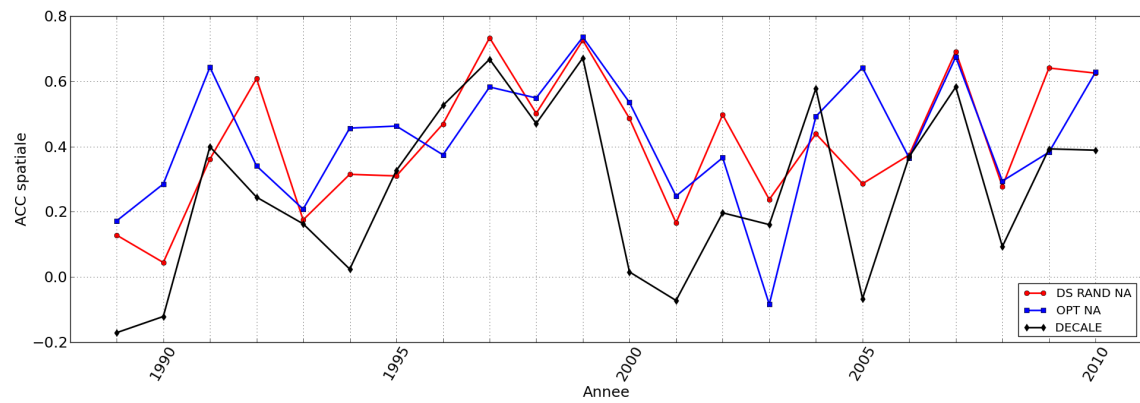
A l'inverse, dans le cas des précipitations sur les tropiques (figure (c)) les ACC spatiales sont dégradées pour 17 années sur les 22 de la période. Pour la T2m sur les tropiques, l'ACC spatiale est meilleure avec les perturbations optimales pour 14 années sur les 22 mais le mACC obtenu avec DS RAND NA est supérieur à celui de OPT NA. L'examen des corrections moyennes de température, pour les années où les perturbations optimales dégradent les scores d'ACC de T2m sur les tropiques, ne montre pas de différence particulière avec les autres années de la période de réanalyse. Dès lors, comment expliquer que les scores soient autant dégradés par l'introduction de ces perturbations optimales ? En ce qui concerne les précipitations tropicales, les résultats suggèrent un effet systématique des perturbations optimales sur les prévisions, comme si les erreurs de tendance initiale du modèle en mode libre sur les tropiques étaient opposées à celles que fait le modèle lorsqu'il est rappelé vers les anomalies d'ERA-Interim. Des expériences de sensibilité aux perturbations ou non des différentes variables (humidité, température et vorticit ) seraient n cessaires pour mieux comprendre ces r sultats. Ces exp riences n cessitent de calculer pour chaque test de sensibilit  une int gration de r f rence et les ensembles de pr vision saisonni re, et n'ont pas  t  impl ment es dans le cadre de cette th se.

En attendant, nous avons voulu tester l'hypoth se selon laquelle les perturbations obtenues *via* le nudging d'anomalie pour une ann e donn e n' taient en fait pas adapt es aux erreurs que fait le mod le lorsqu'il est en pr vision libre cette m me ann e. Pour ce faire, un ensemble de 15 membres avec des perturbations d cal es a  t  impl ment , en tirant les perturbations uniquement dans celles obtenues le m me mois mais l'ann e suivante dans la p riode de r analyse (en prenant 1989 comme ann e suivante pour 2010). Les r sultats obtenus en termes d'ACC spatiale sont montr s dans la figure 5.8 par la courbe noire. Sur les r gions temp r es, les scores de mACC pour Z500 et la T2m sont significativement moins bons qu'avec l'ensemble OPT, mais l g rement meilleurs que DS RAND. Les scores de temp rature sur les tropiques et la r gion Ni o 3.4 sont significativement moins bons que tous les autres ensembles  tudi s, donc encore d grad s par rapport   l'ensemble OPT. Pour les pr cipitations tropicales, le mACC de cet ensemble aux perturbations d cal es est significativement meilleur que OPT, mais toujours significativement moins bon que DS RAND et INI. Toutefois, si on regarde ann e par ann e l'ACC spatiale des pr cipitations tropicales, cet ensemble est meilleur que OPT seulement pour 12 ann es des 22 de la p riode de r analyse  tudi e.

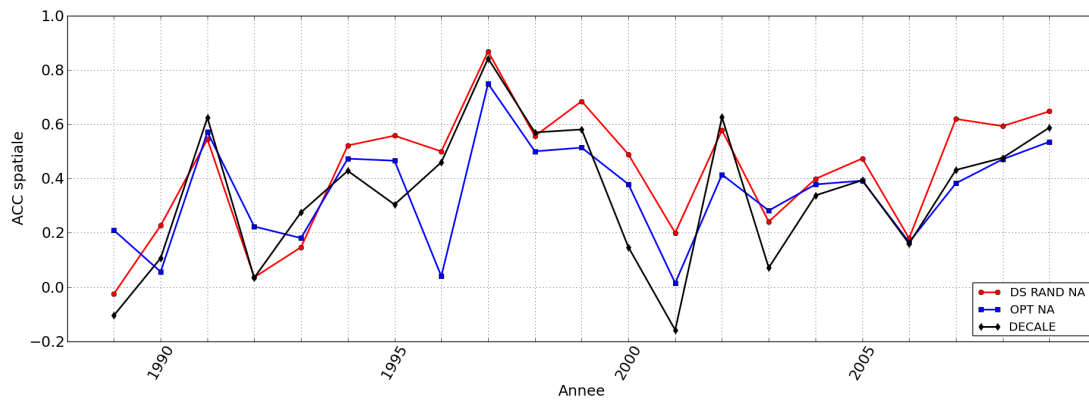
Ces r sultats sont int ressants dans la mesure o  ils montrent que tirer les perturbations dans un m me mois en nudging d'anomalie (donc perturber le mod le en privil giant les corrections de basse fr quence) d grade les pr visions sur les tropiques en hiver par rapport au tirage al atoire haute fr quence. Le tirage de perturbations optimales permet d'att nuer cette d gradation dans le cas de la temp rature de surface, mais pour les pr cipitations il n'y a aucune information utile une ann e donn e dans les perturbations obtenues en nudging d'anomalie. Les am liorations de score obtenues pour les pr cipitations tropicales avec la dynamique stochastique al atoire sont uniquement dues   la meilleure prise en compte des incertitudes du mod le. Cette conclusion est coh rente avec les scores obtenus pour les ensembles DS MENS et MMENS pour ce param tre.



(a) Z500 HN



(b) T2m Tropiques



(c) Précipitations Tropiques

FIGURE 5.8 – ACC spatiale des prévisions saisonnières de (a) Z500 sur l'hémisphère Nord (30°N à 75°N), (b) T2m sur les tropiques, et (c) précipitations sur les tropiques pour les saisons DJF de 1989 à 2010, avec perturbations issues du nudging d'anomalie. Les ensembles comparés sont DS RAND NA (en rouge), OPT NA (en bleu) et l'ensemble avec perturbations décalées d'un an (en noir).

5.1.4 Discussion

Les résultats montrés dans l'article et la comparaison des méthodes nudging d'anomalie et nudging itéré sont intéressants à plusieurs titres. Il est important de noter qu'il n'y a aucune certitude *a priori* d'améliorer considérablement les scores en tirant les perturbations parmi les rappels aux données de référence du mois en cours de prévision. C'est d'autant plus vrai en tenant compte de la répartition entre variabilité interannuelle et variabilité intra-annuelle montrée dans le chapitre précédent : la part de la variance interannuelle étant relativement faible, on pourrait s'attendre à ce que le simple fait de prendre les perturbations du mois courant dans le désordre donne des résultats très similaires à ceux obtenus par la méthode de perturbations aléatoire.

Dans les résultats de cette partie, la dispersion des prévisions avec perturbations optimales n'est que très peu réduite par rapport aux perturbations aléatoires. L'effet dispersif des perturbations, dû à la part de la haute fréquence dans la variabilité des termes des corrections, est toujours présent dans les corrections optimales.

Cependant, les résultats obtenus sur les latitudes tempérées montrent qu'il y a une information utile à l'échelle du mois dans ces termes de perturbations, que ce soient ceux obtenus par le nudging d'anomalie ou le nudging itéré. L'objet de la suite de ce chapitre va être d'explorer des pistes pour extraire cette information utile dans l'état du modèle en cours de prévision, afin d'aller chercher des perturbations appropriées.

Au vu des résultats obtenus avec les perturbations issues du nudging d'anomalie, nous ne nous intéressons par la suite qu'au cas du nudging itéré. Ceci permet également d'étudier le conditionnement des perturbations sur une période de rétro-prévision plus longue (1979–2010).

5.2 Conditionnement en fonction des températures de surface de l'océan

Cette partie présente une méthode de conditionnement des perturbations en fonction de l'état de l'océan au cours de la saison à prévoir.

5.2.1 Principe de base

5.2.1.1 Motivations

L'importance de l'océan en prévision saisonnière n'est plus à démontrer. L'inertie thermique de l'océan étant plus forte, son évolution est plus lente et il garde une « mémoire » plus grande des conditions initiales rapidement oubliées par l'atmosphère. Dans le chapitre introductif, nous avons rappelé l'importance de l'ENSO comme principal mode de variabilité du climat global à l'échelle interannuelle. Le choix est fait ici de se focaliser sur la région du Pacifique tropical pour le conditionnement des perturbations.

5.2.1.2 Implémentation

La durée d’un phénomène ENSO permet de conditionner les perturbations à l’échelle mensuelle. La région Niño 3.4 est choisie pour classer les mois de la période de réanalyse en trois catégories, selon les terciles de température moyenne de la surface de l’océan du mois précédent. Ces trois catégories sont appelées de manière abusive « Niña », « neutre » et « Niño ». Pour un mois M d’une année y des rétro-prévisions, on obtient donc trois populations $\{\delta\mathbf{X}\}_{\text{Niña}}$, $\{\delta\mathbf{X}\}_{\text{neutre}}$ et $\{\delta\mathbf{X}\}_{\text{Niño}}$ contenant les perturbations estimées au cours du mois M des autres années de la période de rétro-prévision, et séparées suivant la catégorie à laquelle appartient la TSO moyenne sur Niño 3.4 du mois $M - 1$ de ces années. Ainsi, les perturbations estimées en février 1998 font partie de la population $\{\delta\mathbf{X}\}_{\text{Niño}}$ (sauf pendant la rétro-prévision de février 1998), car la TSO de janvier 1998 était supérieure au deuxième tercile de la période de réanalyse.

En cours de prévision, une perturbation $\delta\mathbf{X}$ est tirée toutes les 6 heures parmi la population déterminée par la TSO moyenne du mois précédent sur la région Niño 3.4. Ce calcul est bien sûr fait indépendamment pour chaque membre, laissant aux différents membres la possibilité d’être dans des terciles de TSO différents et donc d’être perturbés différemment.

La figure 5.9 montre les covariances des perturbations de température, fonction de courant et humidité spécifique avec les anomalies de TSO sur la région Niño 3.4 le mois précédent, pour les saison DJF et JJA. L’ordre de grandeur de ces covariances est faible par rapport aux corrections moyennes zonales (montrées dans la figure 5.10). Les covariances mettent en évidence une structure cohérente de l’influence des TSO du Pacifique tropical sur les erreurs de tendance initiale du modèle, en particulier pour les champs de fonction de courant. On remarque une anti-symétrie par rapport à l’équateur des corrections de fonction de courant. Les structures des champs de corrections varient entre les deux saisons. Pour l’humidité, les covariances dans les niveaux les plus bas du modèle sont plus fortes en JJA qu’en DJF.

5.2.2 Résultats obtenus

Ce paragraphe détaille l’évaluation des prévisions saisonnières utilisant la méthode de perturbations décrite plus haut. Deux ensembles de 15 membres ont été calculés pour les saisons DJF et JJA, et sont notés DS N34.

5.2.2.1 Prévisions déterministes

Les scores de mACC de DS N34 pour la saison DJF sont montrés dans le tableau 5.4. En comparant les deux colonnes du tableau correspondantes, on constate que l’ensemble DS N34 est significativement moins bon que DS RAND pour les températures de surface sur la région Niño 3.4 et les tropiques, et que le mACC des précipitations tropicales est également dégradé. Sur les régions tempérées, les scores de mACC de DS N34 sont légèrement meilleurs que DS RAND, mais les écarts ne sont pas significatifs.

Le tableau 5.5 montre les scores de mACC de DS N34 dans la troisième colonne pour la saison JJA. Contrairement à l’hiver boréal, les scores de mACC de T2m sur les régions tropicales ne sont pas dégradés par rapport à DS RAND, celui de T2m sur les tropiques étant même significativement meilleur que DS RAND et équivalent à CM5 INI. Par contre, le mACC

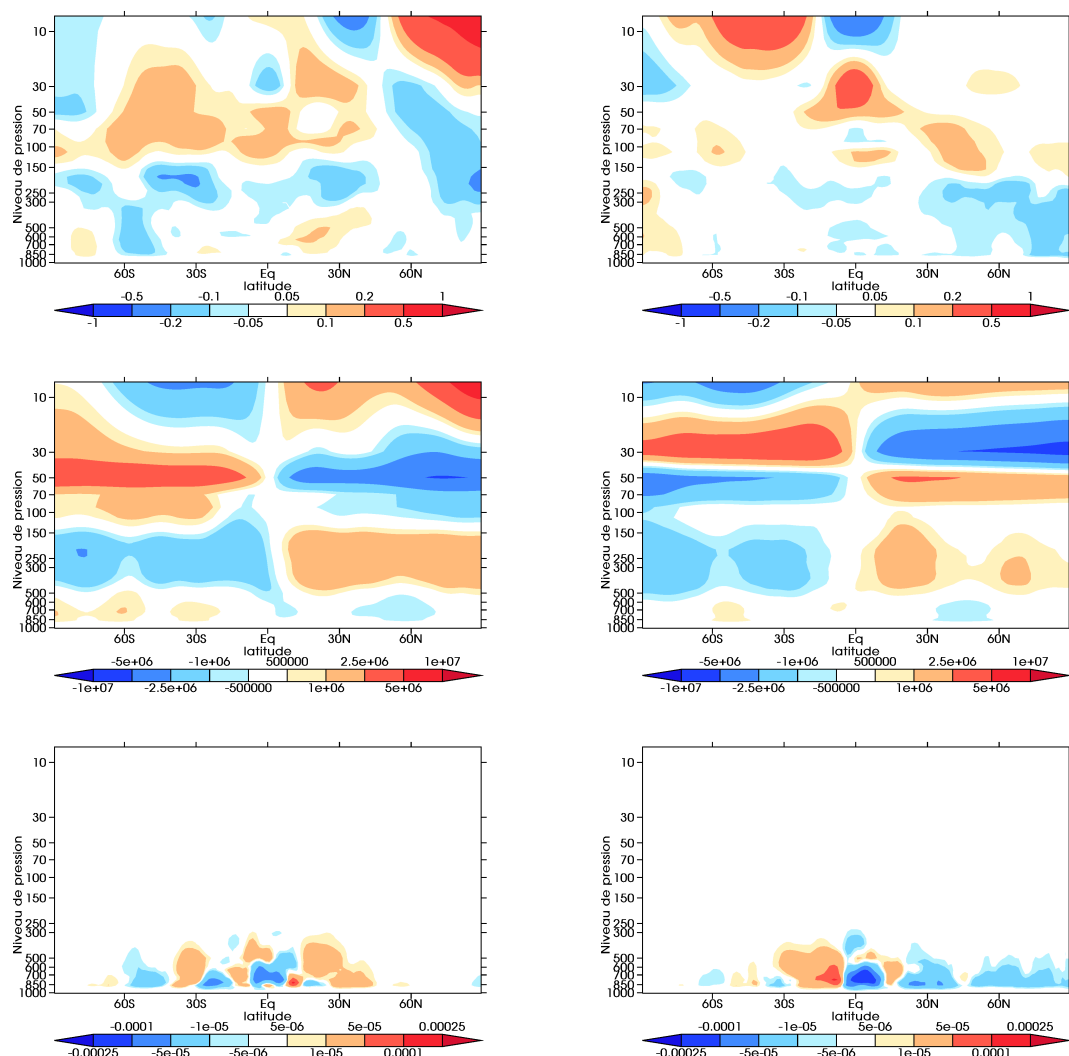


FIGURE 5.9 – Covariances des perturbations moyennes zonales de température, fonction de courant et humidité spécifique (de haut en bas) avec les anomalies de TSO du mois précédent sur la région Niño 3.4, pour les saisons DJF (à gauche) et JJA (à droite). Les échelles sont celles de perturbations (respectivement K/jour, $\text{m}^2\text{s}^{-1}/\text{jour}$ et $\text{kg(eau)}/\text{kg(air)}/\text{jour}$).

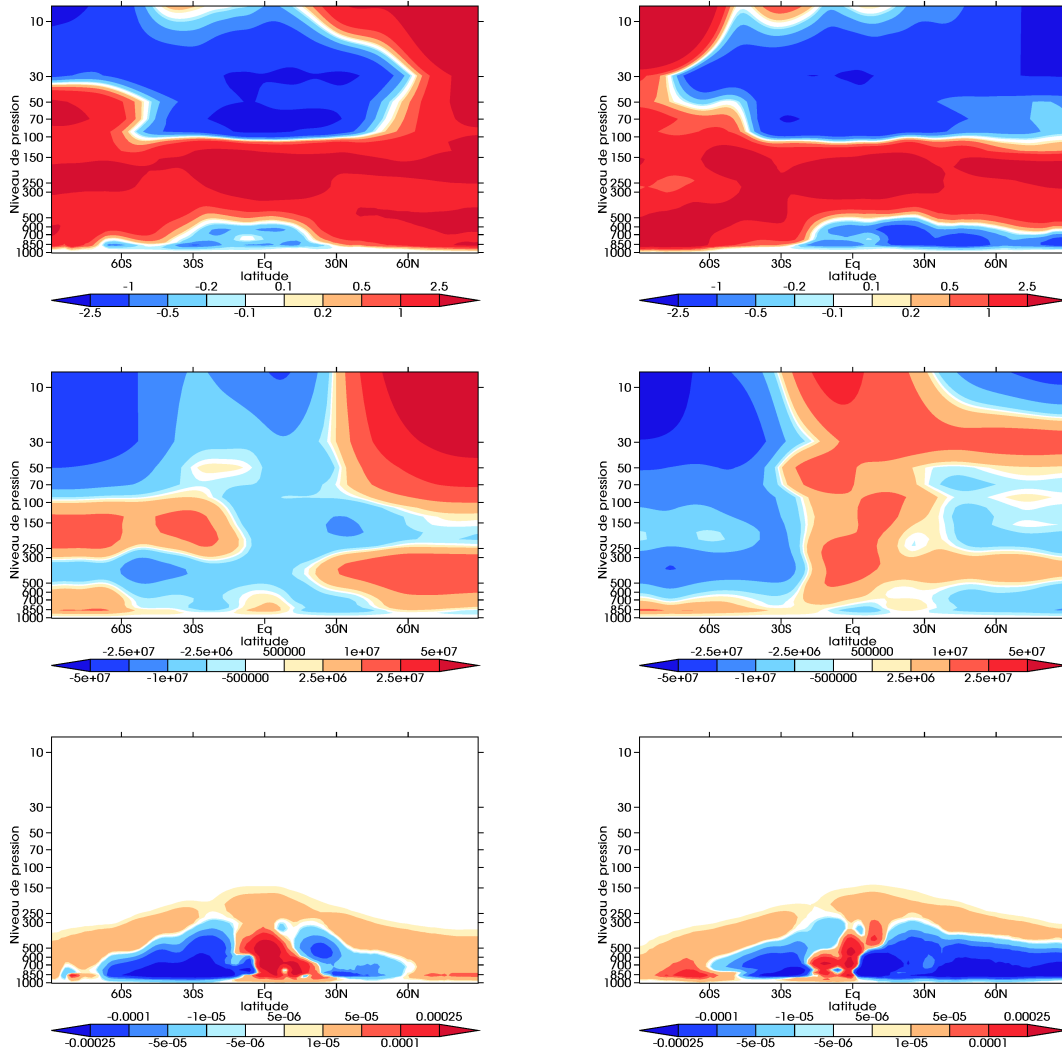


FIGURE 5.10 – Perturbations moyennes zonales de température, fonction de courant et humidité spécifique (de haut en bas) pour les saisons DJF (à gauche) et JJA (à droite). Les échelles sont celles de perturbations (respectivement K/jour, $\text{m}^2\text{s}^{-1}/\text{jour}$ et $\text{kg(eau)}/\text{kg(air)}/\text{jour}$).

Paramètre	Région	CM5 INI	DS RAND	DS N34	DS Ψ 18	DS Ψ 10
T2m	Niño 3.4	0.917	0.911	<i>0.905</i>	0.926^a	0.925
T2m	Tropiques	0.527	0.534	<i>0.520^b</i>	0.534	0.534
T2m	HN	0.25	0.23	0.26	0.26	0.25
Précip.	Tropiques	0.553	0.563	0.551	0.534	<i>0.530</i>
Z500	HN	0.21	0.25	0.28^c	0.24	0.22
Z500	HS	0.28	0.25	0.27	<i>0.19</i>	0.25

^a CM5 INI, DS RAND, DS N34

^b DS RAND, DS Ψ 18, DS Ψ 10

^c CM5 INI, DS Ψ 18, DS Ψ 10

TABLEAU 5.4 – Scores moyens de mACC pour 9 membres parmi 15 (1500 tirages), pour la saison d’hiver DJF et différents ensembles avec dynamique stochastique. Pour les meilleurs et moins bons ensembles, on signale si les écarts de scores avec les autres ensembles sont significatifs par des caractères gras et italiques, respectivement. Lorsque ces écarts sont significatifs pour trois des quatre autres ensembles, on précise lesquels.

Paramètre	Région	CM5 INI	DS RAND	DS N34	DS Ψ 18	DS Ψ 10
T2m	Niño 3.4	0.845	0.826	0.828	0.762	<i>0.755^a</i>
T2m	Tropiques	0.509	<i>0.497</i>	0.506	0.504	0.510
T2m	HN	0.204	0.192	0.198	0.207	0.208
Précip.	Tropiques	0.422	0.410	0.402	0.399	0.403
Z500	HN	0.14	0.10	<i>0.07^b</i>	0.15^c	0.10
Z500	HS	0.29^d	0.23	<i>0.18</i>	0.23	0.24

^a CM5 INI, DS RAND, DS N34

^b CM5 INI, DS Ψ 18, DS Ψ 10

^c DS RAND, DS N34, DS Ψ 10

^d DS RAND, DS N34, DS Ψ 18

TABLEAU 5.5 – Idem tableau 5.4, pour la saison JJA.

des précipitations sur les tropiques est significativement moins bon que ces deux ensembles, de même que le mACC de Z500 sur l’hémisphère Sud. Le score de Z500 sur l’hémisphère Nord est quant à lui significativement dégradé par rapport à CM5 INI et est inclus *in extremis* dans la fourchette des scores de DS RAND. Dans l’ensemble les scores de mACC obtenus avec DS N34 sont inférieurs à ceux de DS RAND, hormis les scores de Z500 sur les régions tempérées en DJF.

En termes de rapport erreur-dispersion, la figure 5.11 montre qu’il y a peu de différences entre les ensembles DS RAND et DS N34 dans l’évolution de l’erreur quadratique moyenne et de la dispersion avec le délai de prévision pendant l’hiver boréal. Les différences sur la région Niño 3.4 apparaissent surtout au cours des deux derniers mois où l’erreur quadratique moyenne de DS RAND est très légèrement inférieure à celle de DS N34 pour la T2m.

5.2.2.2 Prévisions probabilistes

Les scores de RPSS de l’ensemble DS N34 pour la saison DJF peuvent être comparés à ceux de DS RAND dans le tableau 5.6. Les scores sont légèrement dégradés sur les tropiques

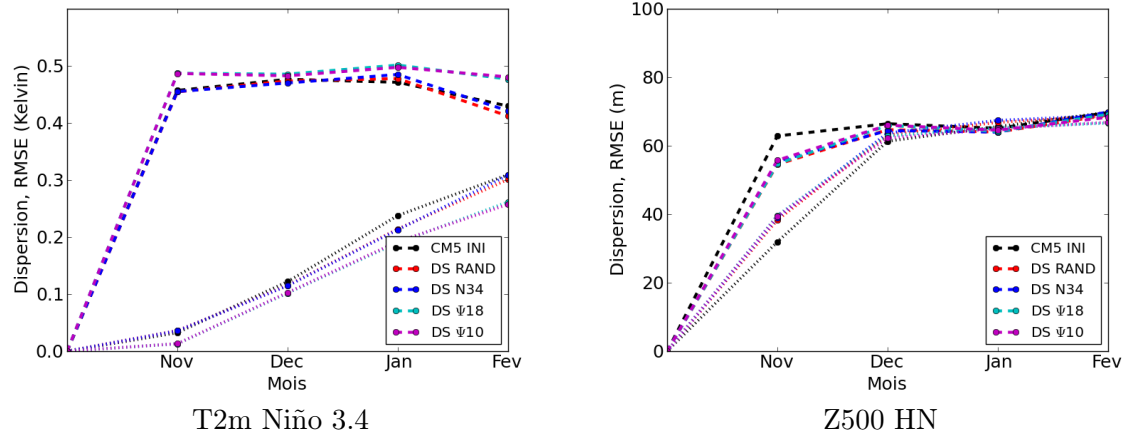


FIGURE 5.11 – Evolution moyenne entre novembre et février du RMSE et de la dispersion des ensembles CM5 INI, DS RAND et ensembles avec conditionnement des perturbations, pour les prévisions de T2m sur la région Niño 3.4 et de Z500 sur l'hémisphère Nord (de 30°N à 75°N).

Paramètre	Ensemble Région	DS RAND			DS N34		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.298	0.908	0.610	0.298	0.900	0.602
Précipitations	Tropiques	0.465	0.513	0.048	0.468	0.513	0.046
Z500	HN	0.408	0.436	0.028	0.415	0.447	0.031
T2m	HN	0.465	0.486	0.021	0.462	0.488	0.026
Paramètre	Ensemble Région	DS Ψ_{18}			DS Ψ_{10}		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.266	0.925	0.659	0.273	0.931	0.658
Précipitations	Tropiques	0.464	0.508	0.044	0.460	0.503	0.043
Z500	HN	0.402	0.424	0.023	0.399	0.422	0.023
T2m	HN	0.436	0.458	0.023	0.434	0.465	0.031

TABEAU 5.6 – Scores de RPSS et décomposition en termes de fiabilité (Rel, divisé ici par l'incertitude Unc) et résolution (Res) pour les ensembles avec conditionnement des perturbations pour DJF 1979–2010.

Paramètre	Ensemble Région	DS RAND			DS N34		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.398	0.762	0.363	0.392	0.768	0.376
Précipitations	Tropiques	0.504	0.532	0.028	0.508	0.528	0.021
Z500	HN	0.383	0.385	0.002	0.385	0.380	-0.005
T2m	HN	0.404	0.404	0.001	0.403	0.406	0.003
Paramètre	Ensemble Région	DS $\Psi 18$			DS $\Psi 10$		
		Rel/Unc	Res/Unc	RPSS	Rel/Unc	Res/Unc	RPSS
T2m	Niño 3.4	0.484	0.753	0.268	0.479	0.740	0.261
Précipitations	Tropiques	0.517	0.537	0.021	0.514	0.537	0.022
Z500	HN	0.397	0.402	0.005	0.411	0.406	-0.005
T2m	HN	0.413	0.413	0.001	0.418	0.417	-0.001

TABLEAU 5.7 – Idem tableau 5.6, pour les expériences avec conditionnement des perturbations pour JJA 1979–2010.

et améliorés sur les latitudes tempérées, conformément aux résultats obtenus pour les scores déterministes. Sur la région Niño 3.4, les prévisions de T2m avec DS N34 ont une moins bonne résolution que les prévisions purement aléatoires. Pour le Z500 sur l'hémisphère Nord, le conditionnement des perturbations par les TSO permet d'améliorer assez nettement la résolution de l'ensemble, mais la fiabilité des prévisions s'en trouve dégradée.

Les prévisions probabilistes de JJA sont évaluées à l'aide du RPSS dans le tableau 5.7. Le RPSS des prévisions de T2m de l'ensemble DS N34 sur la région Niño 3.4 augmente grâce à l'amélioration de la fiabilité et de la résolution par rapport à DS RAND. A l'inverse, ces deux composantes sont dégradées pour les précipitations sur les tropiques et Z500 sur l'hémisphère Nord, tandis que les prévisions de T2m sur l'hémisphère Nord sont équivalentes.

5.2.2.3 Discussion : limites de la méthode

Au final, l'évaluation des prévisions de l'ensemble DS N34 montre que cette méthode de conditionnement des perturbations dégrade les scores de prévision sur les tropiques, et améliore les scores sur les régions tempérées en DJF par rapport au tirage aléatoire des perturbations. Sur la saison JJA, l'ensemble DS N34 est généralement moins bon sur les régions tempérées et légèrement meilleur que DS RAND sur les régions tropicales.

Plusieurs limites à cette méthode peuvent expliquer l'absence d'amélioration des rétro-prévisions. Tout d'abord, les membres d'une prévision évoluent rarement vers des terciles différents de température de surface sur la région Niño 3.4, hormis les cas où la température de surface des états initiaux est proche du premier ou du deuxième tercile de la climatologie. Ces cas sont des cas où le signal de l'ENSO est relativement faible, par rapport à des événements marqués comme ceux de 1982–1983 ou 1997–1998. L'effet du conditionnement ainsi conçu est plus de restreindre la population des perturbations possibles du modèle une saison donnée que de corriger le modèle de manière différente. En témoignent les covariances entre corrections et température sur la région Niño 3.4 qui restent faibles par rapport aux corrections moyennes sur la période de réanalyse.

Qui plus est, la méthode est peu adaptée au cas de la saison d'été (JJA) en raison de la variabilité de l'ENSO. Souvent, l'apparition et la cessation d'un phénomène El Niño ou La Niña

ont lieu au cours des mois d'été. Les populations de corrections d'erreur de tendance initiale vont donc contenir des corrections correspondant à une phase d'amplification d'une anomalie de TSO sur la région Niño 3.4 tout comme celles correspondant à une phase d'atténuation. Or on peut raisonnablement supposer que ces corrections ne seront pas du même acabit.

Ces considérations nous ont mené à envisager un autre type de classification des corrections de tendance initiale du modèle, en conditionnant les perturbations en fonction de classes du champ global de fonction de courant. Cette méthode n'est pas complètement orthogonale à une classification en fonction de l'ENSO, puisque la signature de l'ENSO sur la circulation atmosphérique peut transparaître dans plusieurs analogues de Ψ , au vu de l'importance du phénomène dans la variabilité interannuelle du climat.

5.3 Classification par la fonction de courant

Cette partie présente les méthodes de conditionnement des perturbations en fonction d'analogues de fonction de courant développées dans le cadre de cette thèse. La mise en œuvre de ces méthodes et les résultats obtenus sont détaillés dans ce qui suit.

5.3.1 Principe et mise en œuvre

Le conditionnement des perturbations par rapport à la fonction de courant est basé sur le principe suivant :

1. Le champ global de fonction de courant du modèle au cours de la période de réanalyse est utilisé pour une classification automatique des séries de Ψ au niveau vertical étudié, de manière à déterminer une partition dans l'espace des EOF des champs de fonction de courant en plusieurs classes.
2. Les centres de ces classes sont les analogues de fonction de courant qui vont servir d'étalons par la suite.
3. Les populations de perturbations sont regroupées en autant de classes, selon la classe du champ de Ψ correspondant.
4. En cours de prévision, on détermine de quel étalon le champ courant de Ψ est le plus proche, et on tire les perturbations dans le groupe de perturbations correspondant.

Dans cette thèse, le conditionnement en fonction de Ψ est étudié pour deux niveaux verticaux différents. Les spécificités de chacun de ces conditionnements sont précisées dans les paragraphes correspondants. La classification porte sur des champs globaux, car en cours de prévision les champs de Ψ sont des champs spectraux. On évite ainsi d'ajouter des conversions de champs spectraux en point de grille, afin de ne pas ralentir l'intégration du modèle.

5.3.2 Fonction de courant à 500 hPa

On s'intéresse dans un premier temps à la classification de la fonction de courant à 500 hPa (plus exactement, au niveau 18 du modèle qui correspond à une pression standard de 500 hPa environ).

5.3.2.1 Détails de la méthodologie

Au cours de l'intégration en mode prévision saisonnière nudgée servant à déterminer la population des perturbations de la méthode de nudging itéré, les champs de fonction de courant d'ARPEGE-Climat au niveau 18 sont enregistrés quotidiennement pour chacun des 4 membres.

Afin de procéder à la classification, les champs spectraux sont tronqués à une troncature triangulaire de 21 afin de ne garder que les structures de grande échelle. Ces champs sont lissés dans le temps en calculant des moyennes par fenêtres pentadaires glissantes. Les champs pentadaires ainsi obtenus sont projetés sur leurs n premières EOF, de manière à expliquer 80% de la variance et réduire la dimension de l'espace dans lequel a lieu la classification.

Pour chaque mois de la saison étudiée, on obtient un ensemble de points (coordonnées dans l'espace des EOF) de taille $4 \times n_p \times 32$, où n_p est le nombre de pentades glissantes du mois en question. La classification suit la méthode « k-means ++ », adaptée de la méthode k-means. Cette méthode consiste à itérer plusieurs fois la classification k-means en q classes à partir de q points initiaux tirés aléatoirement parmi l'ensemble de points. On garde la classification qui maximise le rapport entre variance inter-classes et variance intra-classes. La méthode k-means ++ permet en théorie de contourner un écueil fréquent de la classification k-means, qui est la convergence vers un minimum local. En partant d'un grand nombre de jeux de points initiaux différents, on limite les risques de s'enfermer systématiquement dans le même minimum local, et la séparation des différentes classes obtenues est améliorée.

Le choix d'une méthode de classification de type k-means implique de déterminer en amont le nombre de classes recherchées. Un nombre suffisant de classes de champs de Ψ est nécessaire pour décrire des situations météorologiques différentes, mais il faut éviter d'avoir un trop grand nombre de classes. La limite empirique fixée dans le cadre de cette thèse est le nombre de classes tel qu'aucun analogue ne soit défini qu'une seule année de la période de réanalyse. En effet, le tirage des perturbations se fait en validation croisée (de manière à ne jamais tirer une perturbation correspondant à l'année en cours de prévision), aussi veut-on éviter d'avoir des classes non représentées dans les autres années de la période de réanalyse. Dans le cas de la fonction de courant au niveau 18 du modèle, cette méthode a donné un nombre de classes maximum de 8, que nous avons gardé dans la mise en œuvre des prévisions.

L'étape suivante est la répartition des différentes populations de perturbations en fonction de la classe à laquelle appartient le champ de fonction de courant au niveau 18 correspondant, en prenant en compte l'échelle pentadaire. La perturbation définie un jour J dans la rétro-prévision nudgée est classée en fonction de la classe de fonction de courant à laquelle appartient le champ pentadaire de fonction de courant centré sur le jour $J - 3$.

En cours de prévision, les perturbations se font toutes les six heures, mais en changeant de classe de perturbations à un rythme pentadaire. L'initialisation de la prévision (début novembre ou début mai) se fait à partir des états initiaux de l'intégration nudgée faiblement. Les cinq premiers jours, les perturbations des variables pronostiques sont tirées parmi le groupe de perturbations correspondant à l'analogue du champ de fonction de courant de ces états initiaux. Tous les cinq jours, on calcule la moyenne pentadaire du champ de Ψ_{18} et la distance euclidienne de ce champ aux 8 analogues du mois correspondant. Pendant les cinq jours suivants, les perturbations seront tirées parmi la population $\{\delta\mathbf{X}\}_k$ où k désigne la classe correspondant au centre de classe le plus proche.

On procède ainsi pour les 5 premières pentades du mois. La fin du mois est traitée de

manière similaire, en tirant les perturbations la même population $\{\delta\mathbf{X}\}_k$ indépendamment de la longueur du mois. La moyenne des champs de Ψ est calculée sur les jours restants (du 26 à la fin du mois) afin de déterminer la classe dans laquelle on tire les perturbations au début du mois suivant. L'initialisation d'une prévision se fait en déterminant de quel analogue la climatologie du premier mois de prévision sur la période de rétro-prévision est la plus proche.

5.3.2.2 Analogues et corrections

Un exemple des analogues de Ψ_{18} obtenus est donné pour le mois de janvier des prévisions NDJF à la figure 5.12. Les centres d'action des différents analogues de Ψ_{18} sont les plus marqués sur l'océan Pacifique dans l'hémisphère Nord et la région de l'Atlantique Nord. Certaines similitudes entre les différents analogues sont visibles, particulièrement entre les sixième et septième analogues dans le cas illustré ici. Le septième analogue montre des anomalies de fonction de courant beaucoup plus marquées que les autres, mais c'est également l'analogue de la classe la moins peuplée de la période étudiée.

Cette classification a pour but de séparer les populations de corrections d'erreur de tendance initiale en groupes distincts, mais il n'existe aucune certitude avant de procéder à cette classification qu'on obtienne vraiment ce résultat. La figure 5.13 représente la différence entre la moyenne des corrections correspondant à chaque classe du mois de janvier et la moyenne pour toutes les corrections de janvier, pour la fonction de courant au même niveau de 500 hPa. À noter que ces différences sont relativement faibles par rapport aux corrections moyennes (généralement de l'ordre de 10% selon les classes et les zones géographiques). Les différences les plus fortes sont celles liées à l'analogue aux anomalies les plus marquées. Ceci est très certainement dû au plus faible nombre de perturbations dans la population correspondante. Enfin, sur l'Atlantique Nord, certaines différences de perturbations moyennes semblent atténuer les anomalies des analogues correspondants. C'est le cas en particulier des analogues 1, 3, 5 et 8 du mois de janvier comme on peut le voir en comparant les figures 5.13 et 5.12.

La méthode décrite dans cette partie classe les différentes perturbations selon l'analogue de Ψ_{18} auquel appartient l'état du modèle la pentade précédente. On peut se demander si pour une année donnée, les perturbations moyennes en connaissant *a priori* les fréquences des différents analogues dans le mois à venir correspondent aux corrections des erreurs de tendance initiales moyennes du modèle pour ce mois. Pour explorer ce point, on calcule les composites de corrections moyennes en fonction des fréquences d'analogues pour chaque mois de la période DJF 1979–2010 et la corrélation avec les corrections moyennes mensuelles. Les résultats calculés pour chaque point de grille sont montrés dans la figure 5.14. Les corrélations sont plus élevées pour les mois de décembre et janvier que pour février, et on remarque une grande disparité dans les coefficients de corrélation suivant les régions du globe. Toutefois, les corrélations sont positives presque partout, et aucune corrélation n'est fortement négative. Les corrélations les plus fortes sont situées sur les régions tropicales. Ces résultats indiquent que la méthode de perturbation en fonction des analogues de Ψ_{18} est cohérente avec une correction moyenne, tout en présentant des différences en moyenne. Ces calculs de corrélation ne peuvent pas préjuger des résultats : les fréquences de chaque analogue sont ici connues *a priori*, et nous avons vu que les effets non-linéaires des corrections pouvaient avoir une influence sur les prévisions. Une inconnue majeure, avant d'intégrer ces perturbations conditionnées, est l'évolution du comportement du modèle ainsi perturbé par rapport à celui du modèle nudgé faiblement vers des données de référence.

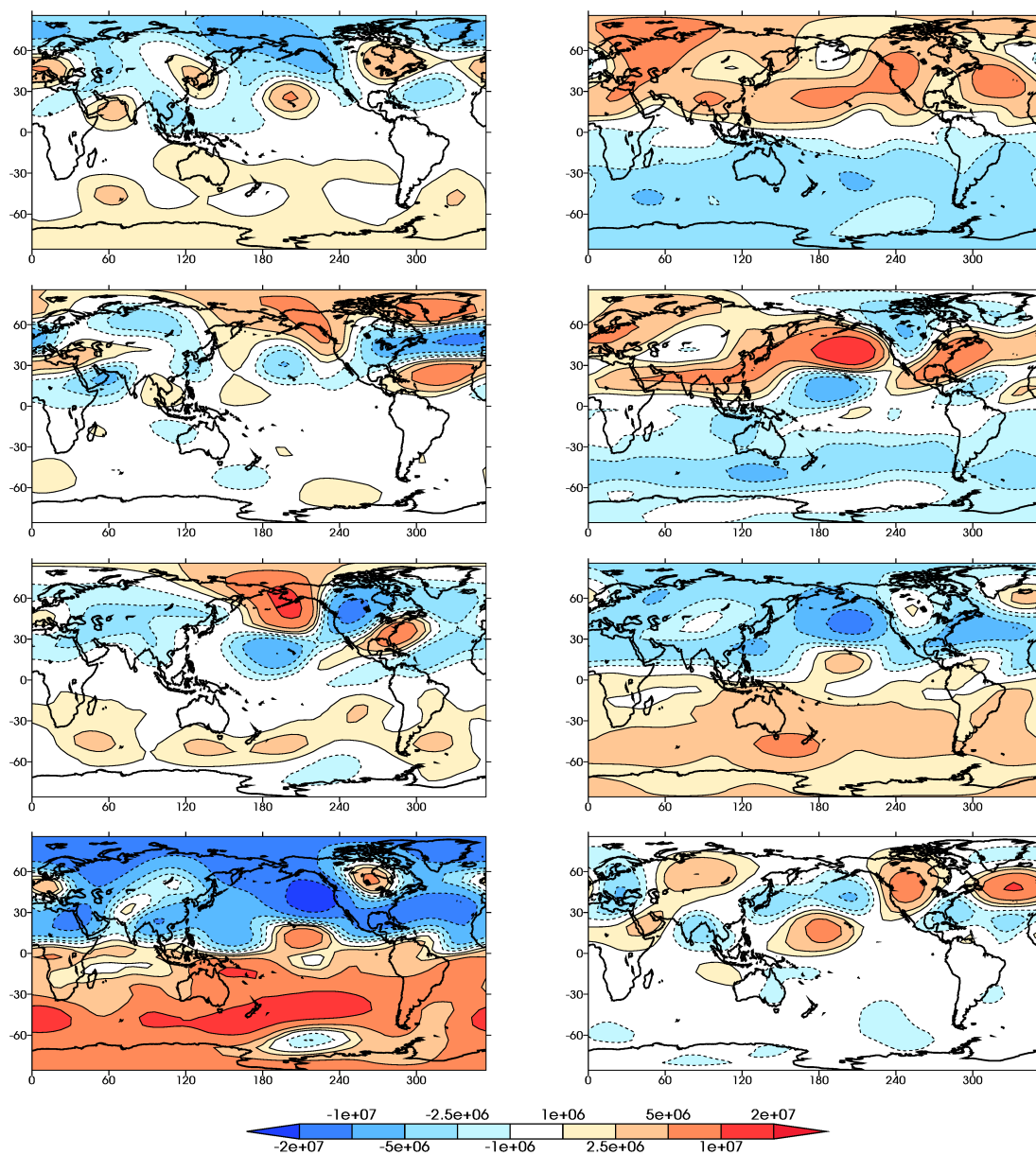


FIGURE 5.12 – Analogues de la fonction de courant à 500 hPa (données des rétro-prévisions nudgées vers ERA-Interim) pour janvier 1980–2011. L'unité est en m^2s^{-1} .

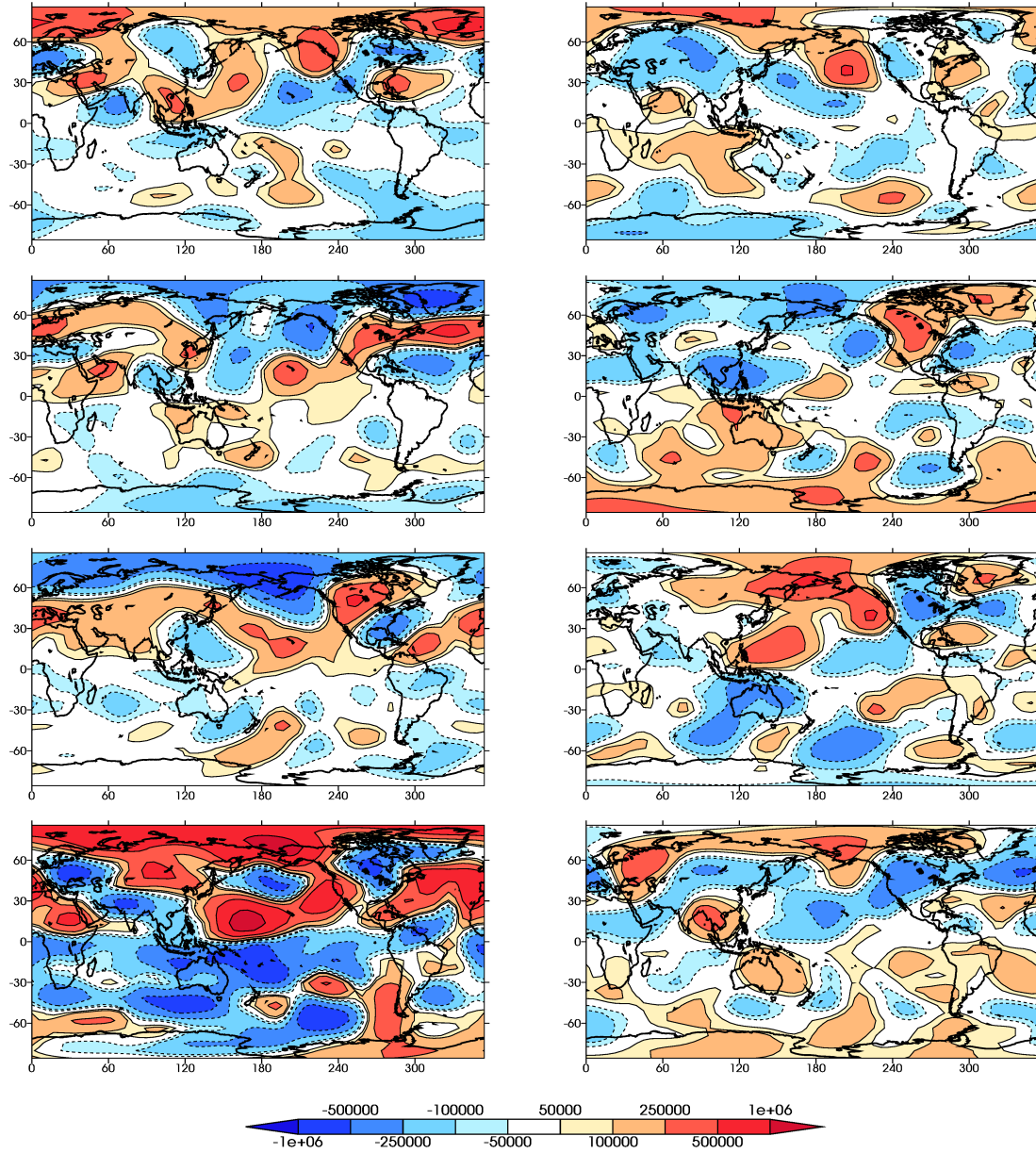


FIGURE 5.13 – Différences moyennes des perturbations de la fonction de courant à 500 hPa pour chaque classe du mois de janvier par rapport à la moyenne de toutes les corrections de janvier. Les classes sont définies à partir des analogues des champs de fonction de courant globaux au même niveau, montrés dans la figure 5.12. L'unité est en $\text{m}^2\text{s}^{-1}/\text{jour}$.

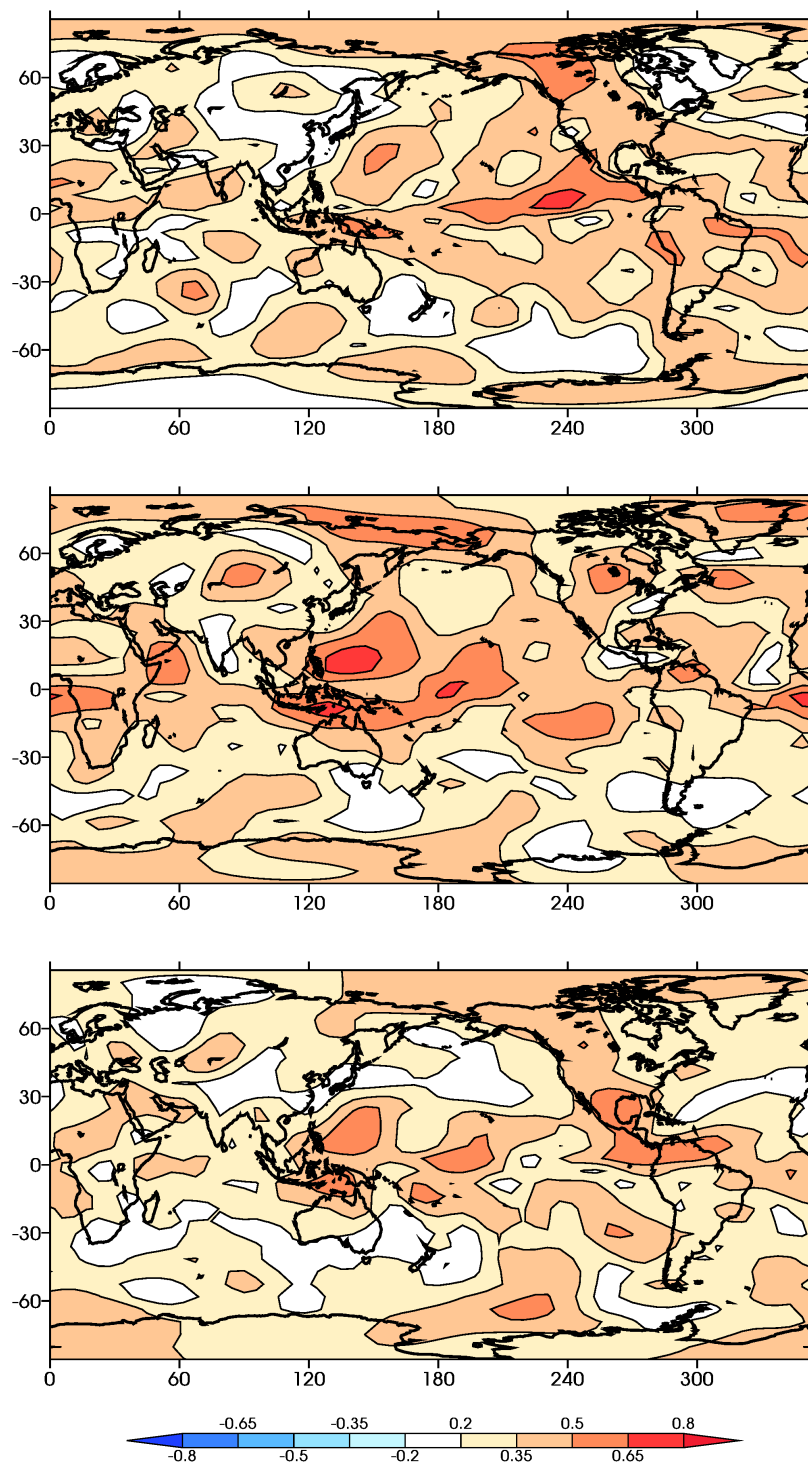


FIGURE 5.14 – Corrélation entre les perturbations moyennes pour un mois donné et les perturbations composites moyennes déduites des fréquences observées des différents analogues de Ψ_{18} , pour la méthode NI et les mois de décembre, janvier et février (de haut en bas).

5.3.2.3 Résultats

Comme précédemment, deux ensembles de 15 membres pour les saisons DJF et JJA, nommés DS Ψ 18, ont été implémentés sur la période de rétro-prévision 1979–2010.

◦ Saison DJF :

Les scores de mACC de la saison DJF sont comparés aux ensembles DS RAND et DS N34 dans le tableau 5.4 (page 162). Sur la région Niño 3.4, les prévisions de T2m sont significativement améliorées. Le mACC de la T2m sur les tropiques est équivalent à DS RAND. Par contre, les prévisions des précipitations tropicales sont significativement dégradées. Les scores de mACC sont également significativement moins bons que ceux de DS RAND sur les régions tempérées pour Z500.

On retrouve ces tendances dans les résultats des prévisions probabilistes, en regardant la décomposition en fiabilité et résolution et le score de RPSS des prévisions de DS Ψ 18 donnés dans le tableau 5.6 (page 163). Le RPSS des prévisions de T2m sur la région Niño 3.4 est nettement meilleur que les ensembles DS RAND et DS N34, tandis que les autres scores étudiés sont équivalents pour T2m sur l'hémisphère Nord et dégradés pour Z500 sur l'hémisphère Nord et les précipitations sur les tropiques.

En termes de rapport erreur/dispersion, la figure 5.11 (page 163) montre que l'erreur quadratique moyenne est plus élevée avec DS Ψ 18 qu'avec les ensembles précédemment étudiés, tandis que la dispersion est plus faible. Ceci n'empêche pas les prévisions saisonnières de cet ensemble d'être meilleures à la fois en termes de mACC et de RPSS que les ensembles DS RAND et DS N34. Dans le cas des prévisions de l'ENSO, l'accroissement du rapport erreur/dispersion n'implique donc pas nécessairement une amélioration des prévisions. Ceci peut être lié au fait que ces prévisions sont déjà d'un bon niveau en termes de scores, et qu'une trop grande dispersion d'ensemble peut dégrader la fiabilité des prévisions. Toutefois, l'ensemble DS Ψ 18 surpasse les ensembles CM5 INI et DS RAND également en matière de résolution pour T2m sur la région Niño 3.4. Sur l'hémisphère Nord aucune différence d'erreur quadratique moyenne ou de dispersion n'est visible pour Z500 avec cette méthode par rapport aux autres méthodes de dynamique stochastique étudiées dans ce chapitre.

◦ Saison JJA :

Le tableau 5.5 (page 162) montre les scores de mACC obtenus pour la saison JJA. À l'inverse de la saison DJF, le mACC sur la région Niño 3.4 est significativement dégradé par rapport aux ensembles CM5 INI, DS RAND et DS N34. Les scores de mACC sur les tropiques sont équivalents à DS N34. Par contre, les scores de Z500 et de T2m sur les régions tempérées sont significativement meilleurs que DS N34, et meilleurs que DS RAND sur l'hémisphère Nord.

Ces résultats sont retrouvés en partie en étudiant les scores de RPSS de l'ensemble DS Ψ 18 dans le tableau 5.7. Le RPSS de la T2m sur la région Niño 3.4 est dégradé, principalement du fait d'une nette augmentation du terme de fiabilité. Les scores de RPSS sur l'hémisphère Nord sont équivalents à ceux de DS RAND, mais la décomposition en fiabilité et résolution varie : la résolution de DS Ψ 18 est légèrement meilleure que DS RAND tandis que la fiabilité est dégradée. Le constat est le même pour les précipitations sur les tropiques, mais le RPSS

est légèrement dégradé par rapport à DS RAND, rejoignant le niveau de DS N34.

5.3.3 Fonction de courant à 200 hPa

La classification des perturbations selon des analogues de fonction de courant à 500 hPa visait à se rapprocher d'une classification en types de temps sur les régions tempérées (Pacifique Nord, Atlantique Nord – Europe). A 200 hPa, la circulation atmosphérique est assez différente de celle à 500 hPa pour qu'on s'intéresse à l'impact sur les prévisions d'un conditionnement des corrections par la fonction de courant à ce niveau.

Les particularités de cette classification et les résultats obtenus pour les saisons DJF et JJA sont présentés dans ce qui suit.

5.3.3.1 Particularités de ces expériences

Contrairement à la fonction de courant à 500 hPa, les champs de fonction de courant au niveau 10 du modèle n'ont pas été stockés au cours de la rétro-prévision saisonnière nudgée servant à établir les populations de perturbations.

Il en résulte quelques adaptations de la méthodologie. Les champs de fonction de courant à 200 hPa sont calculés à partir des données quotidiennes de vent des réanalyses ERA-Interim, puis convertis en données spectrales. Les classes de fonction de courant à 200 hPa sont ensuite calculées à partir des pentades glissantes de ces champs spectraux. Cette méthode implique que les perturbations issues de chacun des 4 membres de la rétro-prévision nudgée correspondant à une date donnée feront partie de la même classe de perturbations.

En cours de prévision, c'est la fonction de courant au niveau 10 du modèle, soit à environ 208 hPa dans l'atmosphère standard, qui sert à déterminer de quelle classe relève la fonction de courant de la pentade précédente, afin de tirer les perturbations dans la population correspondante les cinq jours suivants.

5.3.3.2 Analogues et corrections

Comme dans le cas de la classification de la fonction de courant à 500 hPa, un nombre de huit classes a été retenu pour la classification. Ce nombre permet d'avoir huit populations de perturbations possibles chaque mois, même en validation croisée. La figure 5.15 montre les analogues obtenus pour le mois de janvier des données de réanalyse. Ces résultats nous confortent dans l'objectif de cette expérience : les anomalies de Ψ_{200} sont plus réparties sur l'ensemble du globe que celles de Ψ_{500} , on peut donc espérer obtenir des résultats différents. De fortes anomalies de Ψ sur l'Atlantique Nord et le Pacifique sont encore constatées pour certains analogues.

En termes de corrections, on regarde dans la figure 5.16 les différences entre les moyennes des perturbations au niveau 18 (soit 500 hPa) de chaque classe avec la moyenne sur le mois de janvier. L'échelle de ces différences de perturbations moyennes est la même que pour le cas de la classification selon la fonction de courant au niveau 18 étudiée précédemment (voir la figure 5.13 page 169). Les différences sont proches en intensité et en structure à celles obtenues par

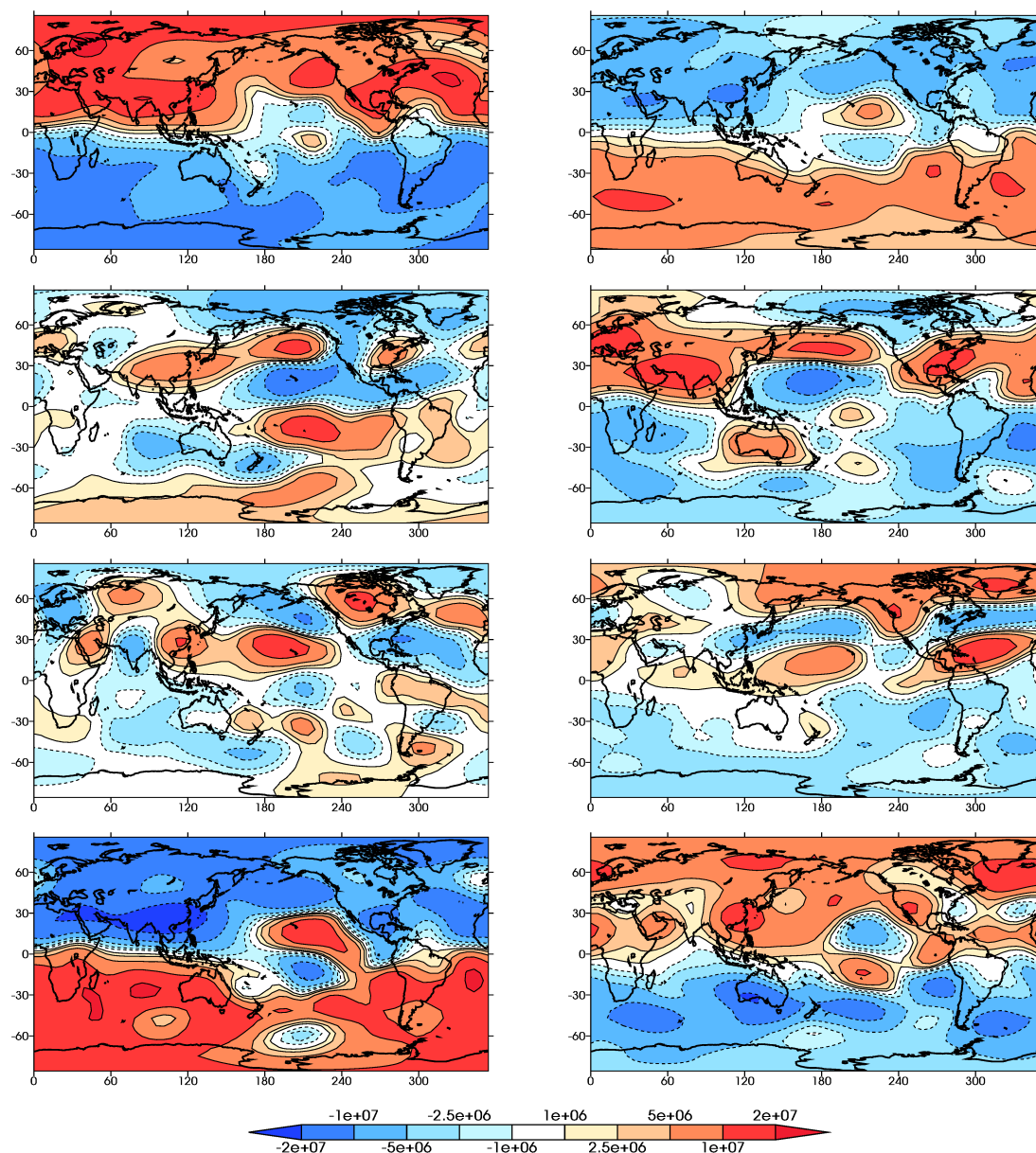


FIGURE 5.15 – Analogues de la fonction de courant à 200 hPa (données globales ERA-Interim) pour janvier 1980–2011.

la méthode précédente. Malgré les différences entre les classes de Ψ aux deux niveaux étudiés, il est possible que les perturbations donnent des résultats similaires.

5.3.3.3 Résultats

La méthode décrite dans cette partie a été testée dans les mêmes conditions que les précédentes, en implémentant des ensembles de 15 membres sur des saisons de 4 mois à partir des conditions initiales de novembre et mai 1979 à 2010 de la première intégration nudgée vers ERA-Interim pour l'atmosphère et les réanalyses NEMOVAR pour l'océan. Ces ensembles sont notés DS Ψ_{10} .

◦ Saison DJF :

Les tableaux 5.4 (page 162) et 5.6 (page 163) donnent les scores de mACC et de RPSS de l'ensemble DS Ψ_{10} pour la saison DJF. DS Ψ_{10} est équivalent à DS Ψ_{18} pour les prévisions de T2m sur les régions tropicales et l'hémisphère Nord en DJF. L'ensemble est équivalent à DS RAND et DS N34 pour Z500 sur l'hémisphère Sud, et légèrement moins bon sur l'hémisphère Sud (de manière non significative). Par contre, les précipitations sur les tropiques sont encore significativement dégradées avec le conditionnement en fonction de Ψ .

Les figures d'erreur quadratique moyenne et de dispersion d'ensemble (figure 5.11 page 163) montrent là encore une très grande similitude entre les ensembles DS Ψ_{10} et DS Ψ_{18} , puisque les courbes sont quasiment superposées.

En termes de prévisions probabilistes, les scores obtenus sont très proches de ceux de DS Ψ_{18} , sauf pour la T2m sur l'hémisphère Nord où une meilleure résolution induit une amélioration du RPSS par rapport à cet ensemble. Par contre, la résolution reste généralement inférieure à celle de DS RAND et DS N34 pour les différents paramètres (mis à part le cas de la T2m sur la région Niño 3.4).

◦ Saison JJA :

Pour la saison JJA, les scores de mACC listés dans le tableau 5.5 (page 162) montrent que les deux conditionnements reposant sur la fonction de courant donnent des résultats très similaires. On observe comme pour DS Ψ_{18} une nette dégradation des scores de mACC de T2m sur la région Niño 3.4. Hormis le score de mACC de Z500 sur l'hémisphère Nord qui est dégradé, les scores de mACC obtenus pour DS Ψ_{10} sont équivalents à ceux de DS Ψ_{18} .

Ces résultats sont confirmés en calculant les scores de RPSS. Les scores et la décomposition en fiabilité et résolution (tableau 5.7, page 164) sont très similaires entre les deux ensembles avec perturbations conditionnées par la fonction de courant. Toutefois, le RPSS de T2m sur la région Niño 3.4 est encore dégradé, en raison d'une moins bonne résolution. La dégradation de la fiabilité des prévisions de Z500 sur l'hémisphère Nord induit un score de RPSS négatif en moyenne sur ces latitudes.

La figure 5.17 représente l'erreur quadratique moyenne et la dispersion d'ensemble des différents ensembles comparés dans ce chapitre pour Z500 sur l'hémisphère Nord et T2m sur la région Niño 3.4 au cours de la saison MJJA. Le rapport erreur/dispersion des prévisions de T2m sur la région Niño 3.4 est très nettement dégradé par le conditionnement des

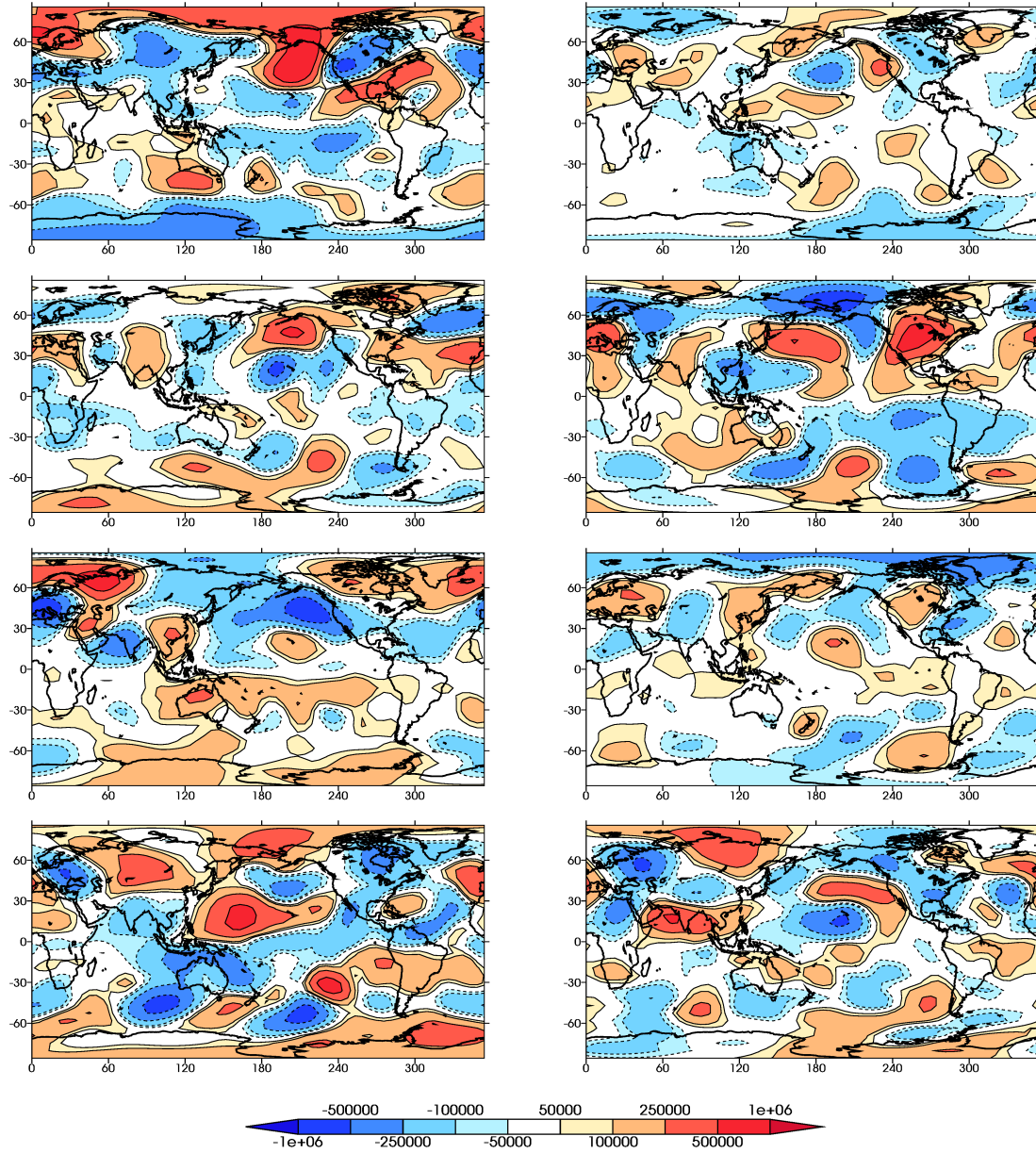


FIGURE 5.16 – Différences moyennes des perturbations de la fonction de courant à 500 hPa pour chaque classe du mois de janvier par rapport à la moyenne de toutes les corrections de janvier. Les classes sont définies à partir des analogues des champs de fonction de courant globaux à 200 hPa des réanalyses ERA-Interim. L'unité est en $\text{m}^2\text{s}^{-1}/\text{jour}$.

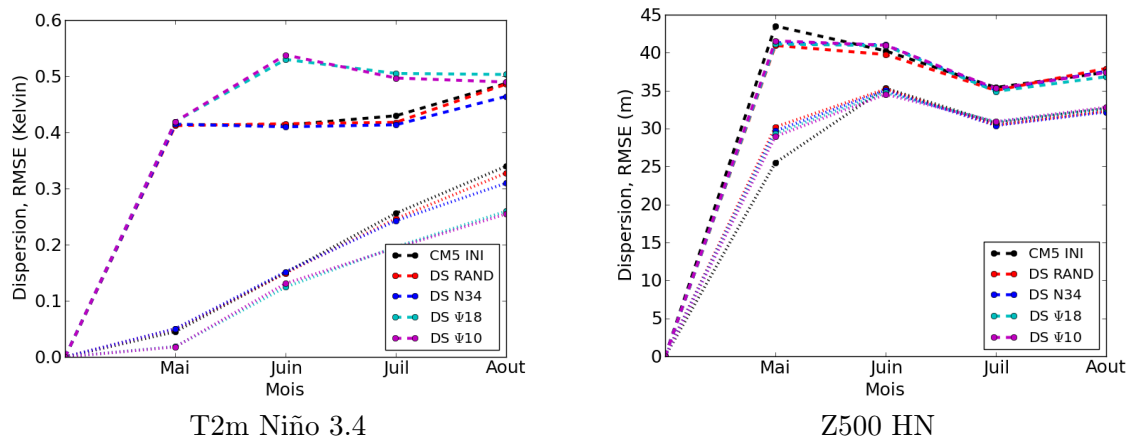


FIGURE 5.17 – Idem figure 5.11, pour les prévisions de mai à août 1979–2010.

perturbations avec la fonction de courant. L'augmentation du RMSE avec les perturbations conditionnées par les classes de fonction de courant apparaît à partir du deuxième mois de prévision. Cette nette dégradation de l'erreur quadratique moyenne est cohérente avec la baisse de fiabilité et la réduction du score de mACC des prévisions.

5.3.4 Interprétation des résultats

Dans l'ensemble, les résultats obtenus avec le conditionnement des perturbations par la fonction de courant sont décevants, dans la mesure où aucune amélioration systématique n'est trouvée par rapport à la méthode de perturbations aléatoires à haute fréquence, et certains scores de la saison JJA sont dégradés.

Plusieurs pistes peuvent être explorées pour expliquer ces résultats. Tout d'abord, la méthode implémentée repose sur l'hypothèse que les analogues de la fonction de courant du modèle nudgé de manière itérée vers ERA-Interim (ou dans le cas de Ψ_{200} , ceux de la fonction de courant calculée à partir des réanalyses ERA-Interim) correspondent bien à l'espace des phases exploré par le modèle en cours de prévision. Ce n'est pas nécessairement le cas pour le modèle sans dynamique stochastique, et le fait d'ajouter des perturbations au modèle peut également changer les caractéristiques de chaque classe de fonction de courant (durée et fréquence d'appartenance) sur la période de rétro-prévision.

Un indicateur de ce problème est le manque de variabilité à l'échelle pentadaire des analogues auxquels s'apparente le champ pentadaire de Ψ_{200} en cours de prévision pour l'ensemble DS Ψ_{10} . Les figures 5.18 et 5.19 montrent les classes dans lesquelles les perturbations sont tirées au cours des saisons NDJF et MJJA, respectivement, pour sept années représentatives de la période de rétro-prévision et les 15 membres des ensembles. Il est utile de rappeler ici que les numéros des analogues sont indépendants entre les différents mois, chaque mois ayant sa propre classification des perturbations et donc ses propres analogues. Pour NDJF, les différents membres n'explorent que deux ou (plus exceptionnellement) trois classes de perturbations différentes au cours d'un mois, et ces classes sont *a fortiori* les mêmes selon les différentes années montrées. Ce constat est encore plus marqué pour la saison MJJA, où tous les membres de la prévision restent proches d'un seul et même membre tout au long des mois de juillet et août

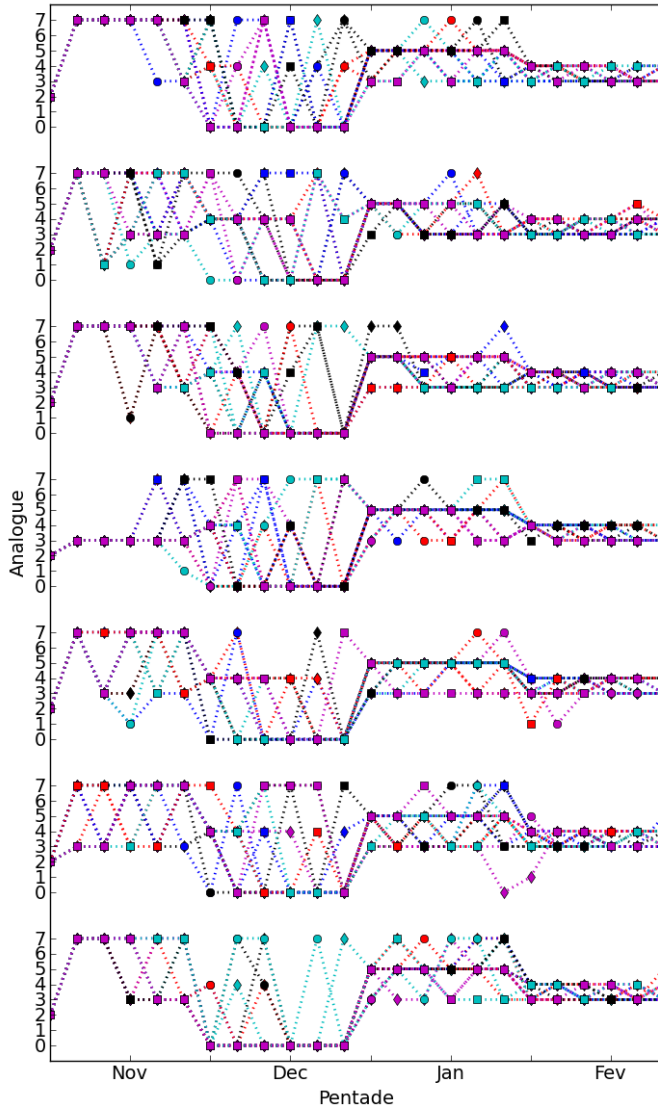


FIGURE 5.18 – Classes d’analogues auxquels appartiennent les champs de Ψ_{200} des 15 membres de prévision au cours de la saison NDJF pour 7 années de la période 1979–2010 (représentatives des différents changements de classes). Sur les 8 analogues définis dans la méthode, on voit que le modèle n’explore qu’un sous-ensemble des classes possibles malgré l’implémentation de 15 membres. Les analogues et leur numérotation changent chaque mois.

sur une des années montrées, et où seuls un ou deux membres s’autorisent des incartades près d’autres analogues pour les autres années de la période de rétro-prévision. La conséquence de ce manque de variabilité est que les perturbations vont être tirées dans un sous-ensemble restreint, qui plus est toujours le même entre les différents membres. C’est une explication possible au manque de dispersion d’ensemble des prévisions sur la région Niño 3.4 constaté sur les figures 5.11 et 5.17.

On pourrait penser que ce manque de variabilité dans le cas de Ψ_{200} est dû au calcul des analogues à partir des données de réanalyses ERA-Interim et non du modèle : il n’en est rien. L’ensemble DS Ψ_{18} présente un comportement similaire en termes de variabilité inter-membres et interannuelle des classes explorées, comme le montre la figure 5.20 page 179. Toutefois, dans ce cas, le nombre de classes explorées par le modèle est généralement plus grand en juillet et août qu’en mai et juin, signe d’un accroissement de la dispersion des membres au fil de la saison prévue. Ceci suggère que dans le cas de la classification en fonction de Ψ_{500} , on peut

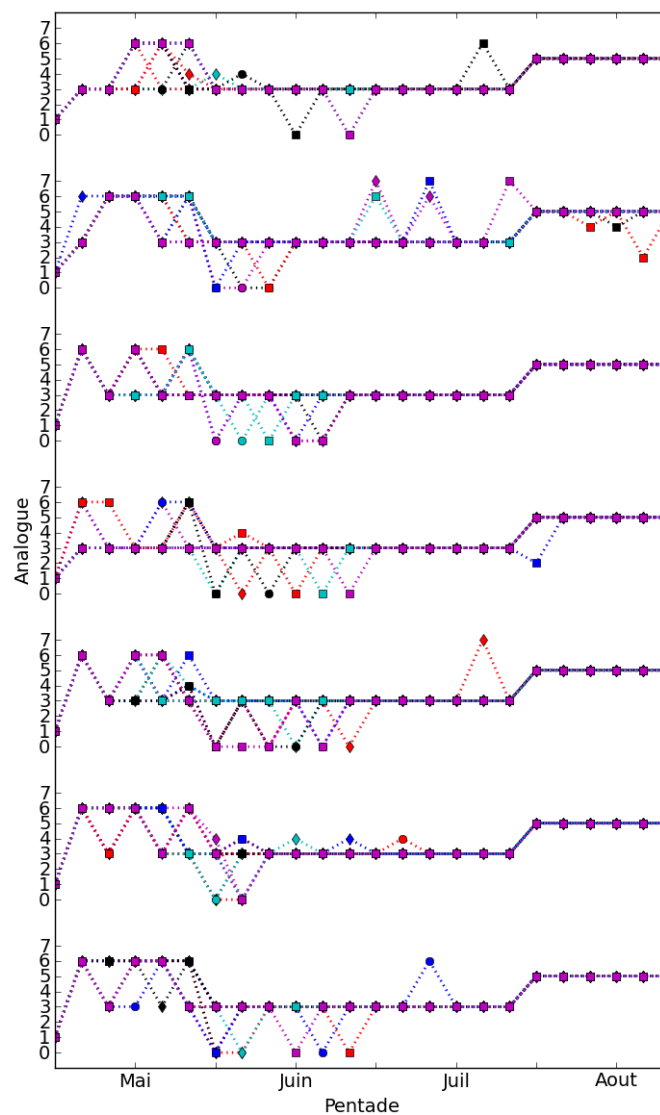


FIGURE 5.19 – Idem figure 5.18, pour la saison MJJA. On a encore choisi 7 années différentes de la période 1979–2010 représentatives des classes explorées par le modèle en cours de prévision. Le constat est très similaire qu’en NDJF, si ce n’est que la faible variabilité entre les membres dans les classes explorées paraît encore plus marquée.

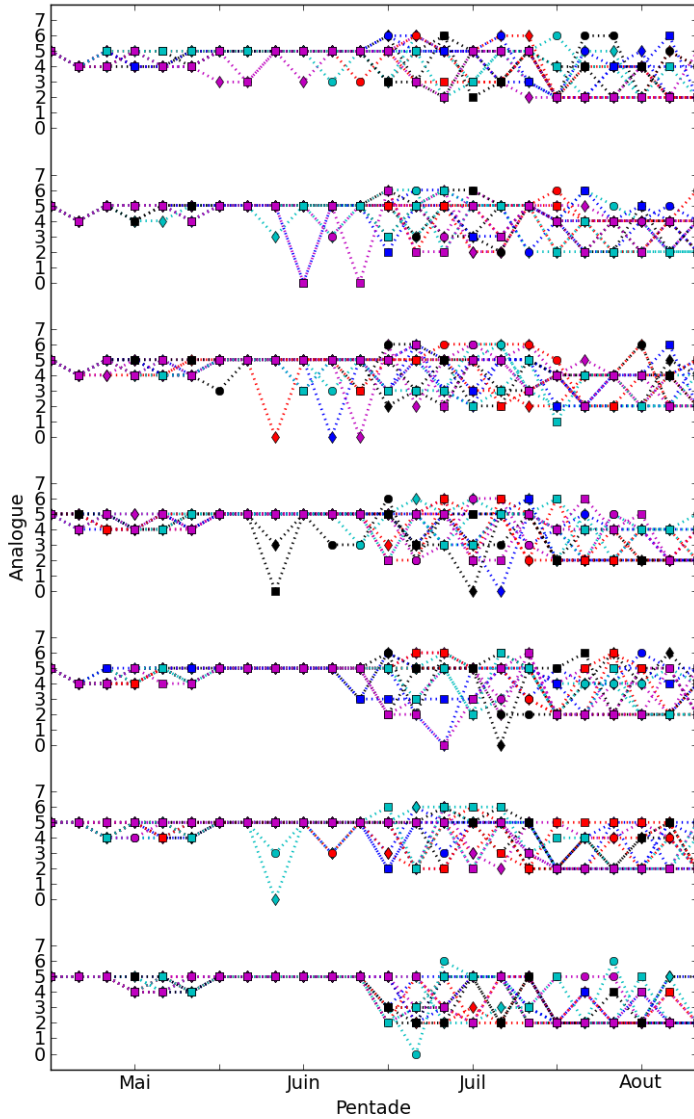


FIGURE 5.20 – Idem figure 5.19, pour l'expérience DS Ψ_{18} et les analogues de Ψ_{500} de la saison MJJA. On a encore choisi 7 années différentes de la période 1979–2010 représentatives des classes explorées par le modèle en cours de prévision. Le constat est très similaire que pour DS Ψ_{10} , avec malgré tout une augmentation du nombre de classes explorées au fur et à mesure de la saison.

espérer augmenter la variabilité des classes explorées en augmentant la dispersion initiale des membres. En effet, la méthode d'initialisation implique de perturber le modèle avec des $\delta\mathbf{X}$ tirés dans la même population pour chaque membre pendant la première pentade du premier mois, et peut certainement être améliorée. Toutefois, l'impact sur les scores reste à étudier.

Une autre piste est celle de la trop grande similitude des corrections des différentes classes obtenues. En effet, les différences moyennes entre les perturbations correspondant aux différentes classes sont assez faibles par rapport à la moyenne des corrections. Ceci est également le cas pour la variance inter-classes des perturbations vis-à-vis de la variance intra-classe. Toutefois, dans le chapitre précédent, nous avons montré qu'un terme entrant en compte pour une part relativement faible de la variance totale des perturbations pouvait contenir les sources d'une amélioration significative de certaines prévisions. Une piste d'amélioration de la méthode serait d'adapter l'algorithme de classification de manière à optimiser la variance inter-classes des corrections plutôt que celle des champs de Ψ .

En conclusion, l'analyse de ces résultats montre que la méthode envisagée n'est pas adaptée au comportement du modèle en mode prévision saisonnière dès lors qu'on ajoute des perturbations en cours d'intégration du modèle. Ces problèmes proviennent de plusieurs sources : d'un manque de séparabilité des corrections en fonction de Ψ d'une part, et d'effets non-linéaires liés à la perturbation haute fréquence des variables pronostiques du modèle ARPEGE-Climat d'autre part, qui font que la méthode de classification déterminée *a priori* ne correspond pas à l'état du modèle en cours de prévision.

5.4 Synthèse

L'objectif de ce chapitre était d'étudier l'impact du conditionnement des perturbations stochastiques du modèle ARPEGE-Climat sur les prévisions saisonnières avec CNRM-CM5.1, en espérant ainsi améliorer la méthode de dynamique stochastique détaillée dans le chapitre 4.

La première partie du chapitre motive cette étude, en montrant qu'une dynamique stochastique « optimale », basée sur le tirage de perturbations calculées dans le mois de la période de réanalyse en cours de prévision, permet d'obtenir des scores de corrélation des anomalies très élevés sur les latitudes tempérées en été comme en hiver. Cette méthode est bien évidemment impossible à mettre en œuvre dans un cadre de prévisions opérationnelles, mais elle indique quelles sont les sorties de prévision qu'on peut espérer améliorer en tirant des perturbations plus adaptées en cours d'intégration du modèle. La comparaison entre les méthodes basées sur le nudging simple itéré et le nudging d'anomalie pour la définition des perturbations nous a incité à nous concentrer dans la suite du chapitre sur la première solution.

Deux méthodes de conditionnement des perturbations ont été implémentées et testées dans cette thèse. La première repose sur le tirage de perturbations en fonction de l'état moyen de l'océan au cours du mois précédent. On a choisi dans ce chapitre de définir pour chaque mois trois populations de perturbations, selon les terciles de température de surface de l'océan Pacifique tropical sur la région Niño 3.4. Les résultats de cette méthode sont contrastés. Si pour la saison DJF, les prévisions sur les régions tempérées sont améliorées par rapport à la dynamique stochastique aléatoire, les scores sur les tropiques sont légèrement dégradés. Pour la saison JJA, on obtient plutôt l'effet inverse. La méthode paraît peu adaptée au cas de la saison d'été boréal, et la séparation des perturbations en trois populations équivalentes en taille est un conditionnement assez peu marqué, comme le suggèrent les covariances des corrections en fonction de ces catégories de TSO sur la région Niño 3.4 montrées dans ce chapitre.

Une deuxième méthode de conditionnement, portant sur la classification automatique de champs de fonction de courant, a été étudiée. Deux jeux de prévision ont été implémentés pour les saisons DJF et JJA, en basant les perturbations sur les analogues de fonction de courant aux niveaux correspondant à 500 hPa puis 200 hPa, respectivement. Les résultats obtenus sont très similaires entre les deux méthodes, à la fois en termes de scores, de dispersion des ensembles et de comportement du modèle en cours d'intégration. Là encore, certaines améliorations obtenues au cours de la saison DJF ne tiennent pas pendant la saison JJA, avec parfois de nettes dégradations des prévisions par rapport aux autres méthodes de perturbation, comme c'est le cas pour la prévision de la T2m sur la région Niño 3.4 en été. Toutefois, la variabilité des populations d'analogues explorées par le modèle ARPEGE-Climat en cours de prévision est trop restreinte. La définition des analogues à partir de données de réanalyse ou

d'une intégration nudgée du modèle couplé sur la période de rétro-prévision n'est donc pas adaptée au comportement du modèle en cours de prévision saisonnière une fois perturbé. Des effets non-linéaires constituent un rempart entre les possibilités de classification *a priori* des perturbations et les améliorations observées avec des perturbations optimales, qui paraît pour l'heure difficile à franchir.

Si ce chapitre constitue une approche à la question du conditionnement des perturbations en vue d'une prévision saisonnière améliorée, il n'est en rien exhaustif. Deux principales solutions de conditionnement ont été explorées, mais beaucoup de pistes sont encore à étudier avant de conclure quant à la possibilité ou non d'un conditionnement efficace des perturbations stochastiques du modèle. Ces pistes sont évoquées dans le chapitre suivant, qui conclut ce travail de thèse.

Chapter Summary

This chapter focuses on state-dependent stochastic perturbation techniques to try and improve seasonal predictions using CNRM-CM5.1 and the stochastic dynamics method detailed in chapter 4.

The first part of the chapter is based on research published in Geophysical Research Letters, where an optimal stochastic dynamics technique shows substantial improvements in ACC scores over mid-latitude regions. This optimal method cannot be implemented for real-time predictions, since it relies on random corrections for initial tendency errors estimated with the NI technique for the current month of the forecast. However, results highlight the regions where improvements can be made, granted that a proper classification of the perturbations derived from the hindcast run can help select the perturbations, so as to correspond to the errors the model tends to make in such a situation. Optimal stochastic dynamics were tested for both DJF and JJA seasons for both NI and NA techniques. Owing to more promising results, the rest of the chapter focuses on the NI technique.

Two perturbation classification techniques were implemented and tested in this thesis. The first relies on classifying the perturbations according to the mean state of the ocean during the preceding month. Results are shown for drawing perturbations in three populations according to Niño 3.4 SST terciles. They differ between seasons and regions of study : in DJF, mid-latitude forecasts are improved with respect to the random stochastic dynamics technique, whereas scores over the tropics are lower. For JJA, opposite results are found. The method is less coherent for JJA since dividing the perturbations according to SST terciles is less efficient during a season where ENSO can be in an onset or decay phase.

A second technique relies on statistical classification of global streamfunction fields. Two state-dependent perturbation techniques were implemented for DJF and JJA, based on a classification of perturbations according to 500 hPa and 200 hPa streamfunction fields, respectively. Results are quite similar between both methods in terms of scores, ensemble spread, and model behavior during the run. Again, improvements found in DJF do not hold for the JJA season. Sometimes, results are significantly worse than with the other stochastic dynamics methods, as in the case of JJA Niño 3.4 2-meter temperature forecasts. A look at the variability of the classes according to which perturbations are drawn highlights some issues in the classification technique. Fewer classes are explored during the run than expected, meaning that once again, the nudged coupled model is not a good approximation for model behavior during seasonal predictions, even in the case of the NI technique.

This chapter paves the way for future work on the issue of state-dependent perturbations of the CNRM-CM5.1 model.

CONCLUSION

CONCLUSIONS ET PERSPECTIVES

Ce travail de thèse s'est intéressé à deux méthodes de prise en compte des incertitudes de modélisation dans les prévisions d'ensemble saisonnières à l'aide de MCG. L'évaluation des prévisions des précipitations du projet ENSEMBLES sur plusieurs régions d'Afrique a permis d'illustrer les avantages de l'approche multi-modèle. Nous nous sommes ensuite focalisés sur la conception et la mise en place d'une méthode appelée dynamique stochastique dans la composante atmosphérique du modèle couplé CNRM-CM5.1. Plusieurs versions de ces perturbations de la dynamique du modèle ARPEGE-Climat ont été testées, et donnent des résultats variant fortement selon la région et le paramètre à prévoir. Après avoir rappelé les principaux résultats et discuté des limites de ces travaux, nous verrons que ceux-ci ouvrent plusieurs perspectives.

6.1 Synthèse générale et discussion

Cette partie reprend les principales conclusions de cette thèse, en les remettant dans le contexte des problématiques générales soulevées en introduction, et en discutant les limites de la démarche scientifique employée.

6.1.1 Prévisions des précipitations africaines : état de l'art des modèles

Les résultats du chapitre 3 confirment l'intérêt de l'approche multi-modèle pour la prise en compte des incertitudes liées à la modélisation. Ils corroborent les conclusions des travaux de Hagedorn et al. (2005) et Alessandri et al. (2011) sur l'approche multi-modèle à l'échelle saisonnière, tout en se focalisant sur un volet moins mis en avant, car souvent problématique : la prévision des précipitations. L'article paru dans *Tellus A* montre que le multi-modèle du projet européen ENSEMBLES apporte une plus-value par rapport à l'utilisation de la climatologie pour déterminer les précipitations moyennes saisonnières sur les régions de l'Afrique de l'ouest en JJA, l'Afrique du sud en DJF, et la Corne de l'Afrique pendant les « short rains » en SON. Les résultats sur la Corne de l'Afrique pendant les « long rains » sont plus mitigés. Un point important de ce travail est la confirmation de l'importance d'une bonne représentation des liens entre TSO et anomalies des précipitations dans les modèles. L'approche multi-modèle est

en cela bénéfique, car elle permet d'atténuer sensiblement les erreurs des modèles individuels.

L'étude des prévisions des précipitations du modèle CNRM-CM5.1, et la comparaison des résultats sur ces régions d'Afrique avec le modèle du CNRM utilisé dans le projet ENSEMBLES, permettent de quantifier les apports et pertes de performances dus aux évolutions du modèle couplé. Les scores déterministes et probabilistes ne sont pas systématiquement meilleurs avec CNRM-CM5.1, mais la représentation du système climatique (liens entre TSO et précipitations, climatologie de l'évolution zonale de la mousson ouest-africaine) est généralement améliorée.

Plusieurs conclusions découlent de ces constats. D'une part, l'approche multi-modèle demeure une technique efficace d'amélioration des prévisions saisonnières, en raison des insuffisances des modèles individuels sur certaines régions pour lesquelles les précipitations sont délicates à prévoir. D'autre part, il faut persévérer dans l'utilisation de modèles dynamiques pour la prévision à l'échéance saisonnière, dans la mesure où les progrès des différentes composantes des MCG sont tangibles à l'échelle de la dizaine d'années.

Nous avons volontairement laissé de côté les approches statistico-dynamiques utilisées pour améliorer les prévisions saisonnières, dont plusieurs ont notamment été développées pour la prévision à l'échelle saisonnière de la mousson africaine, par exemple dans le cadre du projet AMMA. Rappelons ici que ces méthodes peuvent permettre d'accroître significativement les scores par rapport à l'utilisation des sorties brutes d'un ensemble de simulations d'un ou plusieurs MCG. Le lecteur pourra se référer sur ce point aux travaux de Garric et al. (2002), Rogel et al. (2006), Philippon et al. (2010) ou encore Ndiaye et al. (2011).

L'approche multi-modèle n'a d'intérêt que sous certaines conditions : les modèles individuels doivent être comparables du point de vue des performances, et bien calibrés en termes de dispersion. En effet, un modèle sous-dispersif, s'il fait de grosses erreurs de prévision sur une région, tendra à brouiller l'information d'un multi-modèle. Certaines méthodes de calibration visent à pondérer les différents membres du multi-modèle de manière à privilégier l'information provenant des ensembles jugés les plus fiables. Dans ce contexte, la réduction des erreurs des modèles et l'amélioration de la dispersion d'ensemble sont des efforts indispensables à entreprendre pour chaque modèle individuel.

6.1.2 Dynamique stochastique aléatoire

La volonté conjointe de réduire les erreurs du modèle couplé CNRM-CM5.1 et d'améliorer la dispersion d'ensemble des prévisions saisonnières obtenues a conduit au développement de la méthode de dynamique stochastique. Cette technique, décrite dans le chapitre 4, repose sur le tirage aléatoire en cours d'intégration de perturbations additives des variables pronostiques du modèle. Les perturbations correspondent à des corrections d'estimations d'erreurs de tendance initiale de la composante ARPEGE-Climat. La dynamique stochastique permet donc de perturber les champs globaux de température, fonction de courant et humidité spécifique de manière cohérente, puisque les perturbations correspondent à des écarts simultanés de ces différents champs aux données de référence.

Deux méthodes d'estimation des erreurs de tendance initiale ont été étudiées dans le cadre de cette thèse, basées sur le nudging des anomalies de CNRM-CM5.1 vers celles des réanalyses servant de référence (NA), et un nudging itéré vers les réanalyses visant à estimer plus fidèlement l'évolution des erreurs de tendance initiale en fonction du délai de la prévision

Région/Paramètre	NA		NI	
	DJF	JJA	DJF	JJA
T2m Niño 3.4	+	+	—	—
Tropiques	=	+	+	—
Z500 HN	—	—	+	—

TABEAU 6.1 – Résumé de l’impact de la dynamique stochastique aléatoire sur les prévisions saisonnières par rapport aux méthodes de perturbations initiales correspondantes, pour les saisons DJF et JJA et plusieurs paramètres étudiés ; les symboles en gras indiquent une amélioration ou une dégradation significative des scores de mACC, et précipitations et T2m sur les tropiques sont regroupés en une seule catégorie « tropiques », car les résultats sont les mêmes pour ces deux paramètres.

saisonnière (NI). L’apport de la dynamique stochastique aléatoire (consistant à tirer une nouvelle perturbation toutes les six heures en cours d’intégration) a été évalué par rapport à des ensembles partant des mêmes conditions initiales provenant des simulations nudgées, mais perturbés pour chaque membre uniquement à l’instant $t = 0$. Les résultats montrés au chapitre 4 sont contrastés, selon la méthode de définition des perturbations, la saison, la région et le paramètre étudiés. En général, les scores probabilistes confirment les conclusions obtenues avec les scores déterministes.

Le tableau 6.1 résume de manière schématique les résultats obtenus. La méthode NA permet d’améliorer les prévisions sur les régions tropicales, pour les deux saisons DJF et JJA. Ceci se fait au détriment des régions tempérées. Dans le cas de la méthode NI, les perturbations aléatoires permettent de considérablement réduire les biais de Z500 du modèle couplé pour la saison DJF. Ceci se traduit par une amélioration significative de la prévision des anomalies de Z500 pour cette saison, ainsi que des indices NAO et NAM. La méthode NI n’est toutefois pas adaptée à la saison JJA, puisque tous les scores sont dégradés avec l’introduction de perturbations stochastiques aléatoires.

Une étude plus approfondie des perturbations estimées avec la méthode NI a permis d’en savoir plus sur la source des améliorations obtenues en DJF, et d’évaluer la sensibilité des résultats à certains choix d’implémentation. Un autre ensemble perturbé à l’instant initial a été implémenté, en ajoutant à chaque pas de temps aux variables pronostiques une correction moyenne mensuelle, calculée sur les autres années de la période de rétro-prévision. Les résultats obtenus avec cet ensemble (MMENS) montrent que les améliorations (dégradations) des prévisions de Z500 en DJF (JJA) obtenues avec les perturbations aléatoires à haute fréquence sont dues au terme moyen mensuel. L’impact de la fréquence des perturbations a été testé en tirant des perturbations moyennes aléatoires à l’échelle mensuelle. Cet ensemble est plus dispersif mais les résultats sur les régions tempérées sont nettement dégradés. Enfin, une méthode de tirage de séquences de 5 jours de perturbations, qu’on applique ensuite pour 5 jours de prévision, a donné des résultats intéressants. Cette méthode permet d’améliorer encore les prévisions déterministes en DJF (sauf sur la région Niño 3.4) et de retrouver en JJA le niveau des prévisions avec perturbations initiales sur les régions tropicales. Les figures 6.1 et 6.2 regroupent les résultats obtenus pour toutes les versions de la méthode NI en DJF, en termes de mACC et de RPSS, respectivement.

Il peut paraître contradictoire de vouloir à la fois augmenter la dispersion d’ensemble et réduire les erreurs du modèle. Ce point mérite donc une discussion plus détaillée. La réduction des erreurs du modèle est un effet moyen qu’on espère obtenir par une succession de corrections d’erreurs systématiques à chaque pas de temps. D’un autre côté, le fait d’appliquer des

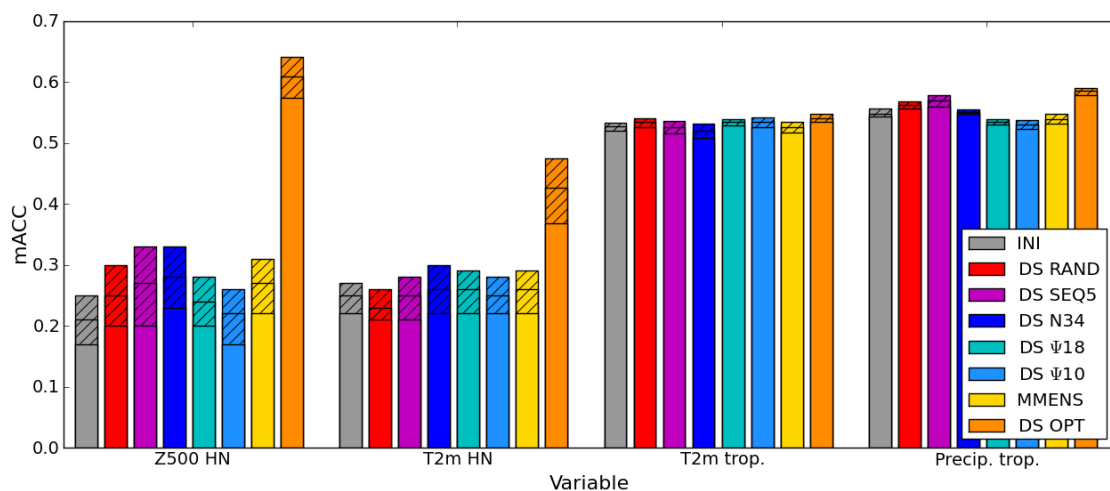


FIGURE 6.1 – Récapitulatif des scores de mACC obtenus pour DJF 1979-2010 et différents paramètres et régions d'étude, pour les ensembles provenant de la méthode NI.

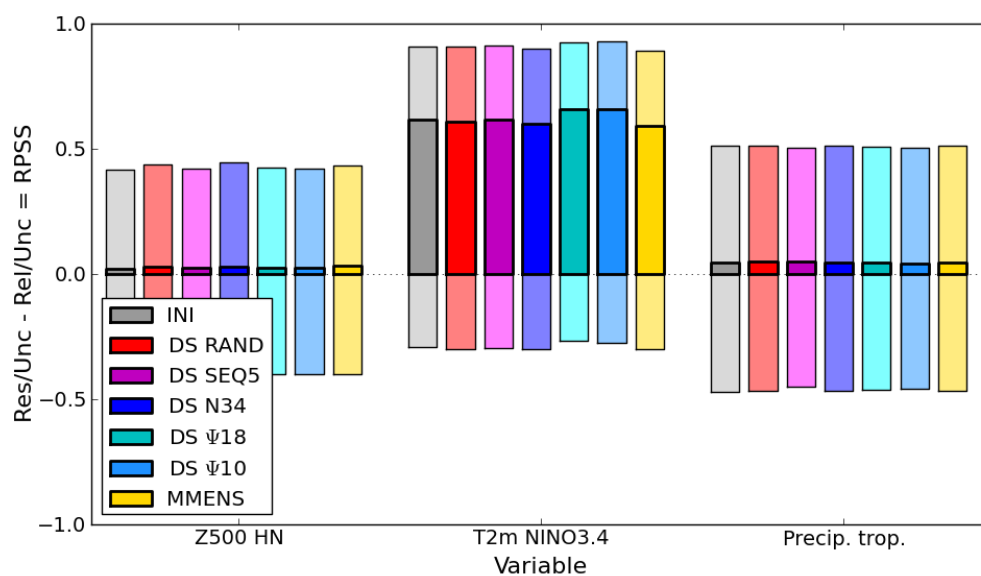


FIGURE 6.2 – Idem figure 6.1, pour les scores de RPSS. La décomposition en résolution (terme positif) et fiabilité (multipliée par -1, terme négatif) est montrée en filigrane, et les scores de RPSS en gras.

corrections évaluées à un instant donné au cours d'une rétro-prévision nudgée, et différentes pour chaque membre de l'ensemble de rétro-prévision libre, doit permettre d'accroître la dispersion d'ensemble du modèle en introduisant un terme de perturbation à haute fréquence. Les deux effets dépendent de la même méthode mais reposent sur des échelles de temps de perturbation différentes.

La décomposition des termes moyens quadratiques des perturbations (figure 4.5 pour la méthode NI) a permis d'illustrer ce point. La part du terme quadratique moyen dans la décomposition est non négligeable. Elle montre que les perturbations peuvent avoir un effet de correction d'erreurs systématiques, tel qu'observé par Guldberg et al. (2005) avec une méthode de corrections moyennes en cours d'intégration, à condition que ces corrections soient appropriées. Le fait que ce terme soit non nul pour la méthode de nudging d'anomalie montre que cette définition des perturbations n'est pas pertinente dans cette optique. Pour les variables et niveaux examinés, la variance intra-mensuelle explique une grande part du terme moyen quadratique. La méthode de dynamique stochastique aléatoire, avec tirage d'un nouveau jeu de perturbations toutes les six heures, se rapproche donc bien de perturbations stochastiques à haute fréquence.

Le troisième terme est celui de la variance interannuelle. La méthode de dynamique stochastique « optimale », décrite au chapitre 5, montre que si ce terme n'explique que 10 à 20% de la variance totale, il comporte une part d'information capitale pour l'amélioration des prévisions. Il est bien évidemment illusoire de penser atteindre les niveaux de scores obtenus avec les perturbations tirées dans le mois courant de prévision. Ces résultats montrent par contre qu'avoir la séquence exacte des corrections pour un mois donné n'est pas nécessaire. L'idée du dernier chapitre de la thèse était de trouver une certaine cohérence entre les corrections correspondant à un état donné du climat, afin de tendre vers ces améliorations.

6.1.3 Conditionnement des perturbations

L'idée d'un conditionnement de corrections d'un modèle de climat n'est pas neuve, mais nous avons cherché à l'appliquer tout en gardant une part d'aléatoire dans les perturbations de la dynamique du modèle ARPEGE-Climat.

Deux méthodes de conditionnement ont été développées dans cette thèse. La première repose sur le tirage de perturbations en fonction de l'état de l'océan au cours du mois précédent, et a été testée pour les terciles de TSO sur la région Niño 3.4. Les résultats sont opposés entre les saisons JJA et DJF. En hiver boréal, les prévisions sur les régions tempérées sont améliorées, tandis que celles sur les régions tropicales sont légèrement dégradées. Cette première approche au conditionnement se veut simple, mais donne des résultats surprenants. Ainsi, les prévisions de T2m sur la région Niño 3.4 sont par exemple significativement dégradées en DJF, preuve s'il en faut de l'aspect non-local de l'impact des perturbations de la dynamique d'un MCG.

Une deuxième méthode s'appuie sur une classification automatique de champs globaux de fonction de courant. Elle a été implémentée pour deux niveaux verticaux différents, à environ 500 hPa et 200 hPa. Les résultats obtenus sont très similaires entre ces deux méthodes, et hormis sur la région Niño 3.4 où les prévisions sont améliorées en DJF et dégradées en JJA, l'impact sur les scores est équivalent pour les deux saisons. L'étude des classes explorées par le modèle a mis en évidence un manque de variabilité par rapport au modèle nudgé, et il en résulte une sous-dispersion d'ensemble, notamment pour la T2m sur la région Niño 3.4.

6.1.4 Principales limites de l'étude

Une des faiblesses de la méthodologie employée dans cette thèse est l'écart entre le comportement du modèle en prévision saisonnière libre et celui du modèle nudgé, même faiblement, vers des données de réanalyses. Ceci est à l'origine de propriétés surprenantes des perturbations issues de la méthode NA. Le choix des temps caractéristiques de nudging est primordial dans la définition de la méthode. Un rappel trop fort peut engendrer des ajustements numériques incohérents du modèle, tandis que si le rappel est trop faible, celui-ci suit une trajectoire qui peut être trop distante des données des réanalyses. La méthode NI a été implémentée après la méthode NA, en tirant parti des limites constatées. A l'évidence, au vu des résultats, une seule itération du nudging simple permet ensuite de corriger certaines erreurs de tendance initiale du modèle, mais pas toutes. Avec la version 4 du modèle ARPEGE-Climat, Déqué (2007) a montré que sept itérations de nudging, en prenant en compte à l'étape suivante la moitié du terme de nudging moyen de l'étape précédente, permettaient de corriger de manière plus efficace les erreurs systématiques du modèle. Toutefois, les impacts en termes de scores de prévision étaient faibles. Il est évident qu'une telle méthode est lourde à mettre en place avec un modèle de climat couplé.

Il est essentiel de souligner ici que les résultats obtenus dépendent sans doute du modèle de climat utilisé au cours de cette thèse. C'est le cas pour de nombreuses méthodes empiriques développées en sciences du climat, qui donnent des résultats variés selon les modèles, l'échéance et les variables étudiées. Le temps imparti n'a pas permis de tester la dynamique stochastique dans un autre modèle couplé. Toutefois, la méthode a été employée pour ARPEGE-Climat dans différentes configurations couplées et à différentes résolutions horizontales et verticales. Les résultats obtenus sont souvent concordants.

Par ailleurs, la prévision saisonnière est une discipline nécessitant des ressources de calcul considérables. Les délais de calcul pour un ensemble de prévision saisonnière étaient d'environ 15 jours pour la rétro-prévision d'une saison sur l'ensemble de la période 1979–2010. Il fallait donc compter un délai d'un mois pour essayer une idée sur les deux saisons DJF et JJA. Plusieurs aspects de la méthode ont donc parfois été modifiés simultanément entre deux expériences, comme dans le cas des prévisions séquentielles de cinq jours. Des expériences intermédiaires auraient pu permettre de trancher quant à l'importance de chaque modification prise séparément. Nous avons toutefois vu que certaines de ces expériences ont été implémentées, avec des résultats allant parfois à l'encontre de l'intuition.

Un dernier facteur limitant en prévision saisonnière est la taille des ensembles. Selon les paramètres étudiés, une taille d'ensemble de 15 membres peut ne pas être suffisante pour trancher quant à l'utilité d'une méthode aux moyennes latitudes.

Ces considérations sont invoquées pour proposer plusieurs perspectives à ce travail de thèse.

6.2 Perspectives

Cette thèse s'inscrit dans un effort d'améliorer les prévisions saisonnières des modèles développés au CNRM depuis maintenant plus de 20 ans. Elle ouvre plusieurs perspectives à des horizons plus ou moins lointains.

6.2.1 Dans un futur proche

La discussion des conclusions de ce travail montre la nécessité de compléter les études des chapitres 4 et 5. En ce qui concerne la fréquence des perturbations, les résultats obtenus avec les séquences de perturbations de cinq jours invitent à implémenter des corrections moyennes à cette fréquence pour évaluer l'intérêt réel de l'aspect séquentiel des perturbations. Le fait que la dynamique stochastique dégrade les prévisions pour la saison de JJA invite à chercher une méthode plus adaptée à cette saison, tout en gardant les améliorations de DJF.

L'étude du conditionnement des perturbations doit également être poursuivie. Une piste possible est celle des classifications reposant sur des méthodes d'analogues sur des régions restreintes. Ce type de classification serait peut-être plus robuste. La réduction de la dispersion avec les méthodes de conditionnement peut être dû à des populations trop petites. La population des perturbations pourrait être aisément agrandie, en augmentant la taille de l'ensemble de rétro-prévision nudgée dans le cas de la méthode NI, ou encore en utilisant un plus grand nombre d'années de données des réanalyses. D'autres centres que le CEPMMT proposent des réanalyses atmosphériques de qualité comparable sur des périodes plus longues.

Une autre piste serait d'étudier une méthode de nudging en point de grille (et non spectral), de manière à rappeler le modèle uniquement sur les tropiques, ou au contraire sur les régions tempérées. Toutefois, il faut garder à l'esprit que les impacts d'un nudging restreint à une région géographique donnée peuvent se ressentir sur des régions très distantes, et il est parfois difficile d'anticiper les effets non-linéaires des perturbations, comme nous l'avons vu dans les résultats de cette thèse.

L'exploration des pistes proposées nécessite parfois de revoir en profondeur la méthode, et de reprendre les étapes d'intégration nudgée du modèle. Dans ce cas, il paraît intéressant d'étudier ces techniques avec un modèle couplé de plus basse résolution (par exemple T63 pour l'atmosphère), en augmentant la taille des ensembles. Ceci permettrait de tester un grand nombre de paramètres (variables pronostiques perturbées, région nudgée, temps caractéristique de rappel, fréquence de perturbations, méthode de conditionnement) à moindre coût.

La technique de dynamique stochastique se focalise sur la perturbation des variables pronostiques de l'atmosphère. Des recherches sont actuellement en cours dans plusieurs centres afin de développer des techniques stochastiques dans les autres composantes d'un modèle couplé de climat. Ainsi, Williams (2012) montre que la perturbation stochastique des flux d'eau douce et de chaleur entre l'atmosphère et l'océan dans un modèle couplé de climat modifie la profondeur de la couche de mélange océanique, la circulation de Hadley et la variabilité de l'ENSO dans une simulation sur cent ans. Au CEPMMT, des travaux sont en cours pour intégrer une méthode SPPT dans le modèle d'océan NEMO. La mise en place de perturbations stochastiques de paramètres d'un modèle de glace de mer est étudiée à l'Institut Alfred Wegener en Allemagne. Ces efforts confirment une tendance actuelle au développement de modèles couplés formulés de manière à inclure de plus en plus de composantes stochastiques, et de telles pistes pourraient être explorées dans un futur proche avec le modèle couplé développé au CNRM.

Ce travail de thèse s'est déroulé entre deux projets européens portant sur la prévision saisonnière à l'aide de modèles numériques : le projet FP6-ENSEMBLES s'est achevé en décembre 2009, et le projet FP7-SPECS a démarré en novembre 2012. L'impact de méthodes stochastiques sur la prévision saisonnière est une des thématiques abordées dans le projet SPECS,

auquel participe le CNRM-GAME.

6.2.2 A plus long terme

Depuis les premières simulations atmosphériques forcées, la prévision saisonnière numérique a fait des progrès tangibles. Les prévisions sur les régions tempérées restent très souvent problématiques, et le chapitre 3 de ce manuscrit a illustré combien il est encore difficile d'obtenir des prévisions utiles sur certaines régions tropicales. L'accroissement espéré des ressources de calcul permettra à l'avenir d'utiliser des modèles numériques à plus haute résolution spatiale et temporelle pour la prévision saisonnière. Dans le cadre du projet Athena piloté par la NSF (US National Science Foundation), des rétro-prévisions saisonnières atmosphériques forcées ont été implémentées pour des années particulières (comme les années de fort El Niño ou d'activité cyclonique intense), à des résolutions horizontales atmosphériques allant de TL159 (environ 126 km) à TL1279 (environ 16 km) avec le modèle du CEPMMT. Les progrès en termes de climatologie et de variabilité à l'échelle saisonnière sont évidents dès le passage d'une résolution TL159 à TL511 (environ 39 km), à la fois sur les régions tropicales et les plus hautes latitudes. Le passage à une résolution encore plus grande ne permet généralement pas d'améliorer autant les résultats, peut-être parce que certaines approximations des modèles de climat ne sont pas adaptées à de telles résolutions (Jung et al., 2012). Cependant, la comparaison d'une méthode de perturbations stochastiques de type « backscatter » à l'augmentation de la résolution horizontale du modèle couplé du CEPMMT à TL511 montre que la physique stochastique peut permettre une réduction des biais équivalente (Berner et al., 2012).

La technique développée dans cette thèse est une méthode de perturbations stochastiques additives. Des travaux récents sur le modèle Lorenz 96 ont confirmé l'intérêt des méthodes stochastiques multiplicatives pour prendre en compte les incertitudes du modèle à l'échelle du climat (Arnold et al., 2012). La formulation d'une méthode stochastique multiplicative dans ARPEGE-Climat pourrait être un objectif à plus long terme, sans garantie toutefois d'améliorer les prévisions : les écarts entre modèles théoriques et modèles numériques couplés d'une grande complexité sont grands, et le seul moyen d'évaluer complètement une méthode stochastique est de l'implémenter dans ces derniers pour la tester.

Terminons ce mémoire de thèse avec l'idée d'une reformulation des modèles de climat de manière à inclure directement l'aspect stochastique des composantes du système Terre dans les équations des modèles, en utilisant la densité de probabilité comme variable pronostique. Cette solution est suggérée notamment par Palmer (2012). Eprouvée d'un point de vue théorique, elle réclamerait des efforts humains et des heures de calcul considérables. Pour l'heure, les améliorations des paramétrisations déterministes, de la puissance de calcul permettant une plus haute résolution, et des techniques de calibration *a priori* et *a posteriori* des sorties des modèles, permettent encore de progresser avec des modèles numériques classiques. Mais gageons que cette solution paraîtra au fil des ans de plus en plus séduisante.

LISTE DES FIGURES

1.1	Trajectoires dans l'espace des phases solutions du système de Lorenz (1963).	6
1.2	Exemple de carte de synthèse des prévisions probabilistes saisonnières pour les précipitations en JJA 2012 par Météo-France.	8
1.3	Position de la thermocline et des zones de convergence, et sens dominant des courants équatoriaux pendant les phases de l'ENSO.	9
1.4	Régions utilisées pour la définition des indices Niño.	10
1.5	Principaux impacts des phases de l'ENSO sur les conditions climatiques des différentes régions du globe en hiver et en été boréal.	11
1.6	Le pattern NAO pour l'hiver boréal.	12
1.7	Pattern de la phase positive du PNA pour différents mois de l'année.	13
1.8	Schéma représentant le principe et les avantages de la prévision d'ensemble avec les modèles de climat.	23
1.9	Analogie du puits de potentiel expliquant l'intérêt des perturbations stochastiques dans un modèle de climat.	26
1.10	Exemple de structure spatiale et d'évolution temporelle du terme r utilisé dans le schéma SPPT.	28
2.1	Répartition sur le globe terrestre des points de la grille réduite TL127 utilisée dans le modèle ARPEGE-Climat v5.2.	37
2.2	Schéma des différentes composantes de CNRM-CM5.1 et des fréquences de couplage.	42
2.3	Principe de la prévision saisonnière d'ensemble entre 2 et 4 mois d'échéance (cas de la prévision de la saison JJA, pour un ensemble de 5 membres).	44
2.4	Un exemple de calibration quantile-quantile des sorties des membres d'une prévision, en vue d'une prévision probabiliste.	53
2.5	Exemple de classification d'un nuage de points en 4 classes.	57
2.6	Régimes de temps sur la région Atlantique Nord-Europe pour la saison DJF 1979–2010, calculés à partir des données quotidiennes de Z500 des réanalyses ERA-Interim.	58

3.1	GPCC seasonal mean precipitation (mm per day) over the African continent for four three-month seasons, averaged over the 1960–2005 period. Boxes show the regions of interest for the given seasons	68
3.2	Seasonal mean precipitation relative anomaly with reference to the 1960–2005 period for GPCC data and the ENSEMBLES project MME ensemble mean, for spatially averaged West Africa JJA precipitation (a) and southern Africa DJF precipitation (b)	70
3.3	Seasonal mean precipitation relative anomaly with reference to the 1960–2005 period for GPCC data and the ENSEMBLES project MME ensemble mean, for spatially averaged Greater Horn of Africa MAM (a) and SON (b) precipitation	71
3.4	Illustration of the quantile-quantile calibration technique used for probabilistic forecasts	73
3.5	Anomaly correlation coefficients over the African continent for the four seasons of interest. Solid and dashed lines indicate positive and negative values, respectively.	76
3.6	Spread-skill scatter plot for the West African region JJA (a) and southern Africa DJF (b) precipitation over the 1960–2005 period. Each point represents spread and ensemble-mean RMSE skill values for a given season and model	78
3.7	RPSS over West Africa for JJA precipitation - only grid points with positive values are shown	78
3.8	Economic value of ENSEMBLES project single models and MME (thick line) for events E^- (left) and E^+ (right) and different cost-loss ratio values (see text for more details) for predictions of JJA rainfall over West Africa	80
3.9	Same as figure 3.6 but for Greater Horn of Africa MAM (a) and SON (b) precipitation. Note that axis scales are not the same for both figures	81
3.10	Same as figure 3.7 but for Greater Horn of Africa MAM (a) and SON (b) precipitation	82
3.11	Economic value of ENSEMBLES stream 2 single models and MME (thick line) for event E^+ and different cost-loss ratio values (see text for more details) for predictions of SON rainfall over GHA	82
3.12	Same as figure 3.7 but for southern Africa DJF precipitation	83
3.13	Same as figure 3.8 but for predictions of DJF precipitation over SA	84
3.14	Covariance between SSTs and standardized precipitation anomalies over GHA (for MAM, SON) WA (for JJA) et SA (for DJF). The left column shows covariances between ERA SSTs and GPCC precipitation and the right column shows the MME covariances. Values are in Kelvin, solid and dashed lines indicate positive and negative values, respectively.	85
3.15	RPSS du multi-modèle ENSEMBLES calculé sur les terciles des précipitations sur la période JJA 1960–2005.	90

3.16	Idem figure 3.14, pour les anomalies de précipitations sur le Sahel et la Guinée dans les données ERA et GPCP et le multi-modèle ENSEMBLES.	92
3.17	Scores de mACC calculés pour chaque mois pour les prévisions de précipitations du projet ENSEMBLES sur le Sahel et la Guinée, pour JJAS 1960–2005.	94
3.18	Idem figure 3.16, pour la saison JAS.	94
3.19	Biais des précipitations et de vent près de la surface en JJA pour les rétro-prévisions du modèle du CNRM dans ENSEMBLES et de CNRM-CM5.1 avec dynamique stochastique séquentielle.	96
3.20	Diagrammes Hovmöller des précipitations quotidiennes (pentades glissantes) de MJJASO pour GPCP, MF ENS et CM5 INI.	97
3.21	Idem figure 3.16, mais sur la période 1979–2005 pour les données GPCP (précipitations) et ERA-Interim (TSO), et les ensembles MF ENS, CM5 INI et DS SEQ5.	98
3.22	Influence de la taille des ensembles sur l’ACC des prévisions de précipitations. .	100
4.1	Décomposition moyenne des termes quadratiques des perturbations pour NDJF et MJJA sur les tropiques et les zones tempérées des hémisphères Nord et Sud, pour la méthode de nudging d’anomalie.	115
4.2	Biais (en m) sur l’hémisphère Nord du Z500 de la saison DJF des ensembles INI et DS.	117
4.3	Moyennes des perturbations $\delta\mathbf{X}$ pour la prévisions nudgée lancée en mai en fonction du délai de prévision pour différents niveaux et variables nudgées. . . .	124
4.4	Idem figure 4.3, pour la prévision nudgée lancée en novembre pour la saison d’hiver NDJF.	125
4.5	Décomposition moyenne des termes quadratiques des perturbations pour NDJF et MJJA sur les tropiques et les zones tempérées des hémisphères Nord et Sud, pour la méthode de nudging itéré.	126
4.6	Biais (en m) sur l’hémisphère Nord (en haut) et l’hémisphère Sud (en bas) du Z500 de la saison DJF des ensembles CM5 INI et DS RAND.	129
4.7	Biais (en m) sur l’hémisphère Nord (en haut) et l’hémisphère Sud (en bas) du Z500 de la saison JJA des ensembles CM5 INI et DS RAND.	131
4.8	Evolution mois par mois de la dispersion d’ensemble et de l’erreur quadratique moyenne au cours des prévisions de NDJF 1989–2010 de T2m sur la région Niño 3.4 et de Z500 sur l’hémisphère Nord, pour les ensembles INI et DS RAND des méthodes NA et NI.	132
4.9	Evolution mois par mois de la mACC des prévisions de T2m sur la région Niño 3.4 pour NDJF 1989–2010 avec les ensembles INI et DS RAND des méthodes NA et NI.	133
4.10	Idem figure 4.8, pour NDJF 1979–2010 et les ensembles basés sur le nudging itéré pour la définition des perturbations.	137

5.1	DJF NH Z500 mean bias (in meters) for ensembles REF (left) and SD_RANDOM (right).	148
5.2	DJF NH Z500 anomaly correlation coefficient for ensembles REF, SD_RANDOM and SD_OPT.	149
5.3	Evolution of monthly root mean square error (full lines) and ensemble spread (dashed lines) for NH Z500 with forecasts REF, SD_RANDOM and SD_OPT. . .	149
5.4	mACC des ensembles INI, DS RAND et OPT basés sur les méthodes NA et NI pour DJF 1989–2010 et différentes régions et paramètres étudiés.	152
5.5	Evolution mois par mois de la dispersion d'ensemble et du RMSE au cours des prévisions NDJF 1989–2010 de Z500 sur l'hémisphère Nord. Les ensembles étudiés sont INI, DS RAND et OPT pour les méthodes NA et NI.	153
5.6	Boîtes à moustaches et moyenne d'ensemble des indices NAO des prévisions de DJF de 1989 à 2010 avec INI, DS RAND et OPT pour les méthodes NA et NI; indices NAO observés sur la période.	154
5.7	Idem figure 5.4, pour JJA 1989–2010. L'absence de fourchettes pour OPT NI provient du fait que c'est un ensemble de 9 membres seulement.	155
5.8	ACC spatiale des prévisions saisonnières de Z500 sur l'hémisphère Nord, T2m sur les tropiques, et précipitations sur les tropiques pour DJF 1989–2010, pour trois ensembles NA.	157
5.9	Covariances des perturbations moyennes zonales de température, fonction de courant et humidité spécifique avec les anomalies de TSO du mois précédent sur la région Niño 3.4 pour DJF et JJA.	160
5.10	Perturbations moyennes zonales de température, fonction de courant et humidité spécifique pour DJF et JJA.	161
5.11	Evolution moyenne entre novembre et février du RMSE et de la dispersion des ensembles CM5 INI, DS RAND et ensembles avec conditionnement des perturbations, pour les prévisions de T2m sur la région Niño 3.4 et de Z500 sur l'hémisphère Nord (de 30°N à 75°N).	163
5.12	Analogues de la fonction de courant à 500 hPa pour janvier 1980–2011.	168
5.13	Différences moyennes des perturbations de la fonction de courant à 500 hPa pour chaque classe du mois de janvier par rapport à la moyenne de toutes les corrections de janvier.	169
5.14	Corrélation entre les perturbations moyennes pour un mois donné et les perturbations composites moyennes déduites des fréquences observées des différents analogues de Ψ_{18} , pour la méthode NI et les mois de décembre, janvier et février (de haut en bas).	170
5.15	Analogues de la fonction de courant à 200 hPa (données globales ERA-Interim) pour janvier 1980–2011.	173

5.16	Différences moyennes des perturbations de la fonction de courant à 500 hPa pour chaque classe de Psi_{200} du mois de janvier par rapport à la moyenne de toutes les corrections de janvier.	175
5.17	Idem figure 5.11, pour les prévisions de mai à août 1979–2010.	176
5.18	Classes d’analogues auxquels appartiennent les champs de Ψ_{200} des 15 membres de prévision au cours de la saison NDJF pour 7 années de 1979-2010.	177
5.19	Idem figure 5.18, pour la saison MJJA.	178
5.20	Idem figure 5.19, pour l’expérience DS Ψ_{18} et les analogues de Ψ_{500} de la saison MJJA.	179
6.1	Récapitulatif des scores de mACC obtenus pour DJF 1979-2010 et différents paramètres et régions d’étude, pour les ensembles provenant de la méthode NI.	188
6.2	Idem figure 6.1, pour les scores de RPSS.	188

LISTE DES TABLEAUX

2.1	Pression et altitude dans l'atmosphère standard des différents niveaux du modèle ARPEGE-Climat à 31 niveaux verticaux.	38
3.1	Basic characteristics of the main components of the five ENSEMBLES stream 2 models used for seasonal forecasting	67
3.2	mACC values for the regions of study and corresponding seasons of interest over the 1960–2005 period and for three 15 or 16-year sub-periods (*) <i>see text for more details</i>	77
3.3	Single model and MME spatial average RPSS (%), RPS, and reliability - resolution decomposition (values divided by uncertainty) for the different regions and seasons of interest over the 1960–2005 period	80
3.4	Classification of DJF seasons in warm, average and cold years categories for Niño 3.4 SST. (The year corresponds to the month of December)	86
3.5	mACC values for the regions of study and corresponding seasons of interest over the 1960–2005 period for the MME over warm, average and cold DJF SST years	86
3.6	Scores mACC et RPSS sur les zones Guinée et Sahel des modèles individuels et du multi-modèle ENSEMBLES en JJA.	89
3.7	Scores mACC en fonction des terciles de TSO sur la région Niño 3.4 pendant la saison DJF qui suit la saison JJA étudiée, sur le Sahel et la Guinée.	91
3.8	Idem tableau 3.6, pour la saison JAS.	93
3.9	zACC sur les régions de l'Afrique de l'ouest pour le modèle de Météo-France du projet ENSEMBLES et plusieurs ensembles utilisant CNRM-CM5.1.	99
3.10	Scores de RPSS et décomposition en Rel et Res pour les prévisions des précipitations en JJA sur l'Afrique de l'ouest des différentes versions du modèle couplé du CNRM.	101
3.11	Idem tableau 3.9, mais pour les régions d'Afrique du sud en DJF et la Corne de l'Afrique en MAM et SON.	101
3.12	Idem tableau 3.10, pour l'Afrique du sud en DJF.	102
3.13	Idem tableau 3.10, pour la Corne de l'Afrique en MAM et SON.	102

4.1	Valeurs de τ et λ pour les expériences en nudging d'anomalie.	113
4.2	Altitude et pression de l'atmosphère standard pour les niveaux du modèle ARPEGE-Climat évoqués dans les chapitres 4 et 5.	116
4.3	Scores de mACC pour DJF 1989–2010 des ensembles INI et DS.	116
4.4	Scores de RPSS et décomposition en Rel et Res pour les ensembles INI et DS de prévisions saisonnières de DJF 1989–2010.	118
4.5	Durée (en jours) et fréquence moyenne (en %) de chaque régime pour les réanalyses ERA-Interim et les ensembles INI et DS (moyenne d'ensemble et écart-type en italique) pour DJF 1989–2010.	119
4.6	Corrélation des fréquences moyennes prévues par INI et DS avec les fréquences des réanalyses pour chacun des régimes, et score de BSS (voir texte).	120
4.7	Idem tableau 4.3, pour la saison JJA.	120
4.8	Idem tableau 4.4, pour JJA 1989–2010.	120
4.9	Valeurs de τ et λ pour la deuxième étape du nudging itéré (nudging faible). . .	122
4.10	Scores moyens de mACC pour 9 membres parmi 15 pour la saison DJF et différents ensembles avec ou sans dynamique stochastique.	128
4.11	Idem tableau 4.4 mais pour les expériences basées sur le nudging itéré pour DJF 1979–2010.	128
4.12	Idem tableau 4.10, pour la saison JJA.	130
4.13	Idem tableau 4.11, pour la saison JJA 1979–2010.	130
4.14	Scores de corrélation des indices de NAO et NAM pour DJF 1989–2010 des ensembles INI et DS RAND des méthodes NA et NI.	134
4.15	Idem tableau 4.11, pour les expériences MMENS, DS MENS et DS SEQ5 et la saison DJF 1979–2010.	135
4.16	Idem tableau 4.15, pour la saison JJA 1979–2010.	135
4.17	Fourchettes à 5% et 95% des ACC de 9 membres parmi 15 des indices NAO et NAM pour DJF 1979–2010 des différents ensembles basés sur le nudging itéré. .	138
4.18	Durée moyenne et fréquence des quatre régimes de temps sur la région Atlantique Nord - Europe pour DJF 1979–2010 pour ERA-Interim et cinq ensembles de prévisions saisonnières basés sur le nudging itéré; ACC des fréquences de régimes et Brier Skill Score de ces ensembles.	139
5.1	Mean ACC values for REF, SD_RANDOM and SD_OPT. Statistical significance of differences between the SD ensembles and REF are tested using a binomial test for season ACC scores. Bold scores are significantly better than REF at a 95% level.	148

5.2	Reliability, resolution, RPS and RPSS values for ERA-Interim climatology, REF, SD_RANDOM and SD_OPT for NH T2m (land grid points only) and Z500. Bold RPS values indicate scores significantly better than REF at a 95% level using a binomial test for season RPS scores.	150
5.3	Scores de corrélation des indices de NAO et NAM pour DJF 1989–2010 des ensembles OPT.	153
5.4	mACC pour 9 membres parmi 15 pour la saison DJF et différents ensembles avec dynamique stochastique.	162
5.5	Idem tableau 5.4, pour la saison JJA.	162
5.6	Scores de RPSS et décomposition en termes de fiabilité (Rel, divisé ici par l'incertitude Unc) et résolution (Res) pour les ensembles avec conditionnement des perturbations pour DJF 1979–2010.	163
5.7	Idem tableau 5.6, pour les expériences avec conditionnement des perturbations pour JJA 1979–2010.	164
6.1	Résumé de l'impact de la dynamique stochastique aléatoire sur les prévisions saisonnières, pour les saisons DJF et JJA et plusieurs paramètres étudiés. . . .	187

LISTE DES SIMULATIONS

Sont listés ici les différents ensembles de rétro-prévision saisonnière implémentés dans le cadre de cette thèse et discutés dans le manuscrit. La méthode de dynamique stochastique a été appliquée en validation croisée, ce qui signifie que pour chaque mois de rétro-prévision on ne pouvait pas tirer de perturbations calculées au cours du mois correspondant dans la simulation nudgée servant à les définir.

Rétro-prévisions de MJJA/NDJF 1989–2010 basées sur le nudging d’anomalie

Les ensembles suivants comportent 15 membres calculés pour les saisons MJJA et NDJF 1989–2010, à partir de conditions initiales issues d’une intégration du modèle CNRM-CM5.1 dont les anomalies sont rappelées vers celles d’ERA-Interim. Les perturbations $\delta\mathbf{X}$ proviennent de cette même intégration.

INI Perturbations initiales avec un $\delta\mathbf{X}$ tiré aléatoirement pour chaque membre ; noté aussi **INI NA** pour la comparaison avec les méthodes NI.

DS Perturbations aléatoires en tirant un nouveau $\delta\mathbf{X}$ toutes les 6 heures parmi les perturbations correspondant au mois calendaire en cours de prévision ; noté aussi **DS RAND NA**.

OPT NA Perturbations optimales tirées toutes les 6 heures parmi celles correspondant au mois et à l’année en cours de rétro-prévision.

Rétro-prévisions de MJJA/NDJF 1979–2010 basées sur un nudging itéré

Les ensembles suivants comportent 15 membres¹ calculés pour MJJA et NDJF 1979–2010, à partir de conditions initiales issues d’une intégration du modèle CNRM-CM5.1 nudgé directement vers les données de réanalyses ERA-Interim. Les perturbations $\delta\mathbf{X}$ proviennent d’une seconde itération du nudging avec des coefficients de rappel plus faibles.

CM5 INI Perturbations initiales avec un $\delta\mathbf{X}$ tiré aléatoirement pour chaque membre ; noté parfois **INI NI**.

DS RAND Perturbations aléatoires en tirant un nouveau $\delta\mathbf{X}$ toutes les 6 heures parmi les perturbations correspondant au mois calendaire en cours de prévision ; noté parfois **DS RAND NI**.

1. sauf DS OPT pour MJJA qui n’en comporte que 9

LISTE DES SIMULATIONS

- MMENS** Perturbations moyennes mensuelles (moyenne sur les autres années de la période).
- DS MENS** Perturbations aléatoires mensuelles en tirant un $\delta\mathbf{X}$ moyen correspondant aux corrections moyennes du mois calendaire mais d'une autre année de la période de rétro-prévision.
- DS SEQ5** Perturbations aléatoires en tirant tous les cinq jours une séquence de $\delta\mathbf{X}$ obtenue au cours de cinq jours consécutifs de la deuxième intégration nudgée, correspondant au mois calendaire en cours de prévision.
- DS OPT** Perturbations optimales tirées toutes les 6 heures parmi celles correspondant au mois et à l'année en cours de prévision.
- DS N34** Perturbations tirées toutes les 6 heures en fonction du mois en cours de prévision, et du tercile de TSO sur la région Niño 3.4 du mois précédent.
- DS Ψ 18** Perturbations tirées toutes les 6 heures en fonction du mois en cours de prévision et de la classe de fonction de courant au niveau 18 du modèle ARPEGE-Climat auquel appartient le champ moyen de la pentade précédente ; on réévalue la classe tous les 5 jours.
- DS Ψ 10** Idem **DS Ψ 10** mais en fonction de la classe de Ψ du niveau 10 du modèle.

ACRONYMES

ACC	Anomaly Correlation Coefficient, page 50
AMMA	African Monsoon Multidisciplinary Analysis, projet international lancé en 2002 pour améliorer les connaissances et la compréhension de la mousson de l'Afrique de l'ouest et sa variabilité à plusieurs échelles spatiales et temporelles, page 31
ARPEGE-Climat	Action de Recherche Petite Echelle Grande Echelle-Climat, modèle d'atmosphère développé à Météo-France pour les études aux échelles saisonnières à climatiques, page 32
Atl 3	boîte Atlantique 3, définie par les longitudes comprises entre 20°W et 0°E et les latitudes entre 3°N et 3°S, page 97
CASBS	Cellular Automaton Stochastic Backscatter Scheme, méthode de perturbations stochastiques utilisée au CEPMMT, page 29
CEPMMT	Centre Européen de Prévisions Météorologiques à Moyen Terme, organisation intergouvernementale européenne qui fournit des prévisions météorologiques à moyen et long terme ainsi que des ressources de calcul pour la recherche scientifique (en anglais : ECMWF), page 17
CERFACS	Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, centre de recherche en méthodes de simulation numérique avancées co-financé par sept actionnaires dont Météo-France, page 31
CMIP5	Coupled Model Intercomparison Project, cinquième projet d'inter-comparaison des modèles couplés de climat, page 35
CNRM-CM5.1	modèle de climat (version 5.1) développé conjointement par le CNRM et le CERFACS, participant à l'exercice d'intercomparaison des modèles couplés CMIP5, page 31
CNRM-GAME	Centre National de Recherches Météorologiques - Groupe d'études de l'Atmosphère Météorologique, unité de recherche associée de Météo-France et du CNRS, et laboratoire très accueillant, page 19
CPC	Climate Prediction Center, subdivision du NCEP en charge de la compréhension des fluctuations du climat à courte échéance, page 13
DEMETER	DEvelopment of a Multi-model Ensemble system of seasonal-to-inTERannual prediction, projet financé par le cinquième programme cadre (FP5) de la commission européenne, page 19

DJF	saison Décembre-Janvier-Février, page 20
DSP	Dynamical Seasonal Prediction, projet coordonné entre cinq groupes de recherche américains en modélisation du climat à la fin des années 1990, page 19
ECMWF	European Centre for Medium-Range Weather Forecasts, voir CEPMMT, page 17
ENSEMBLES	projet européen du sixième programme cadre (FP6) étudiant la prévision d'ensemble à différentes échelles climatiques, comprenant un volet portant sur la prévision saisonnière, page 20
ENSO	El Niño Southern Oscillation, El Niño-Oscillation Australe, page 9
EOF	Empirical Orthogonal Function, page 55
ERA-40	ECMWF Re-Analysis 40, réanalyse atmosphérique globale du CEPMMT portant sur la période 1957–2001 soit une période de plus de 40 ans, page 45
ERA-Interim	ECMWF Re-Analysis Interim, réanalyse atmosphérique globale du CEPMMT portant sur la période 1989–2009, puis étendue à la période 1979–présent, page 45
EUROSIP	EUROpean Seasonal-to-Interannual Prediction, consortium européen (initialement, inter-continental depuis l'intégration du modèle du NCEP) fournissant des prévisions saisonnières multi-modèles en temps réel, page 19
GELATO	Global Experimental Leads and Ice for Atmosphere and Ocean, modèle de glace de mer développée au CNRM et dont la version 5 est utilisée dans le modèle couplé CNRM-CM5.1, page 40
GIEC	Groupe d'experts Intergouvernemental sur l'Evolution du Climat, page 112
GLACE	Global Land-Atmosphere Coupling Experiment, page 15
GPCC	Global Precipitation Climatology Centre, centre fondé en 1989, régi par le service météorologique allemand (Deutscher Wetterdienst) visant à fournir des données précises de précipitation à l'échelle du globe, page 46
GPCP	Global Precipitation Climatology Project, projet international de provision de données d'analyses de précipitations globales aux échelles quotidiennes à mensuelles, page 46
HN	Hémisphère Nord ; sauf mention du contraire, dans cette thèse, lorsque l'acronyme HN est employé dans le contexte du calcul de scores, il désigne les latitudes entre 30°N et 75°N, page 116
HS	Hémisphère Sud ; comme pour HN, dans le contexte du calcul de scores, l'acronyme désigne les latitudes entre 30°S et 75°S, page 120
IPSL	Institut Pierre-Simon Laplace, institut de recherche en sciences de l'environnement regroupant six laboratoires en région parisienne, page 43
JAS	saison Juillet-Août-Septembre, page 93
JJA	saison Juin-Juillet-Août, page 7
mACC	mean-ACC, page 50

MAM	saison Mars-Avril-Mai, page 64
MCG	Modèle de Circulation Générale, page 18
MJJA	saison Mai-Juin-Juillet-Août, page 96
NA	Nudging d'Anomalie, page 112
NAM	Northern Annular Mode, ou mode annulaire de l'hémisphère Nord, page 14
NAO	North Atlantic Oscillation, principal mode de variabilité du climat sur l'Atlantique Nord et l'Europe, page 11
NASA	National Air and Space Administration, page 45
NCEP	National Centers for Environmental Predictions, le pôle climat du National Weather Service américain qui dépend de la NOAA, page 13
NDJF	saison Novembre-Décembre-Janvier-Février, page 113
NEMO	Nucleus for European Modelling of the Ocean, modèle dont la composante océanique est utilisée dans CNRM-CM5.1, page 40
NEMOVAR	données de réanalyses océaniques produites au CEPMMT à l'aide du modèle d'océan NEMO, page 54
NI	Nudging Itéré, page 112
Niño 3.4	région du Pacifique tropical équatorial comprise entre 175 et 120°W servant à définir un indice de TSO caractéristique des phases de l'ENSO, page 10
NOAA	National Oceanic and Atmospheric Administration, agence fédérale des Etats-Unis en charge des prévisions de l'état de l'océan et de l'atmosphère, www.noaa.gov , page 9
OASIS	coupleur numérique développé au CERFACS dont la version 3 est utilisée dans CNRM-CM5.1, page 41
OMM	Organisation Météorologique Mondiale, page 17
OPA	Océan PArallélisé, composante océanique de NEMO, page 44
PDF	Probability Density Function, ou fonction de densité de probabilité, page 27
PNA	Pacific/North American pattern, mode de variabilité du climat sur l'Amérique du Nord et l'océan Pacifique, page 13
POTENTIALS	Project on Tendency Evaluations using New Techniques to Improve Atmospheric Long-term Simulations, projet de l'Union Européenne initié deux ans après PROVOST, page 20
PRESAO	PREvisions Saisonnières en Afrique de l'Ouest, forum visant à fournir chaque année des prévisions statistiques et dynamiques de la saison de mousson sur la région, page 63
PROVOST	PRediction Of climate Variations On Seasonal to interannual Time-scales, projet de recherche financé par l'Union Européenne de 1997 à 1999, page 19
QBO	Quasi-Biennial Oscillation, mode de variabilité des vents zonaux équatoriaux dans la stratosphère, page 14

RMSE	Root Mean Square Error, ou erreur quadratique moyenne, page 48
RPS	Ranked Probability Score, page 52
RPSS	Ranked Probability Skill Score, page 52
SAM	Southern Annular Mode, ou mode annulaire de l'hémisphère Sud, page 14
SOI	Southern Oscillation Index, page 10
SON	saison Septembre-Octobre-Novembre, page 64
SPECS	Seasonal-to-decadal climate Prediction for the improvement of European Climate Services, projet financé par le septième programme cadre (FP7) de la Commission Européenne depuis fin 2012, page 191
SPPT	Stochastically Perturbed Parametrization Tendencies, méthode de perturbations stochastiques développée au CEPMMT, page 27
SURFEX	SURFace EXternalisée, interface entre la composante d'atmosphère et le schéma de surface continentale du modèle CNRM-CM5.1. SURFEX calcule également les flux au-dessus de l'océan., page 40
T2m	Température à 2 mètres, page 116
TSO	Températures de Surface de l'Océan, page 9
WCRP	World Climate Research Program, page 46
Z500	Géopotential à 500 hPa, page 116
zACC	Transformée en Z de Fisher de la série des ACC spatiales calculées chaque année pour une saison donnée, page 50
ZCIT	Zone de Convergence Inter-Tropicale, page 10
$\delta\mathbf{X}$	Terme de perturbation stochastique correspondant à une correction d'erreur de tendance initiale du modèle ARPEGE-Climat, page 111
Ψ	Fonction de courant; Ψ_n désigne la fonction de courant au niveau n du modèle ARPEGE ou à n hPa des données étudiées (selon les cas), page 165

BIBLIOGRAPHIE

- A. Alessandri, A. Borrelli, A. Navarra, A. Arribas, M. Déqué, P. Rogel, et A. Weisheimer (2011). Evaluation of probabilistic quality and value of the ENSEMBLES multimodel seasonal forecasts : comparison with DEMETER. *Monthly Weather Review*, 139 : 581–607. doi : 10.1175/2010MWR3417.1.
- M. H. P. Ambaum, B. J. Hoskins, et D. B. Stephenson (2001). Arctic Oscillation or North Atlantic Oscillation? *Journal of Climate*, 14 : 3495–3507.
- H. M. Arnold, I. M. Moroz, et T. N. Palmer (2012). Stochastic parametrisations and model uncertainty in the Lorenz '96 system. *Philosophical Transactions of the Royal Society of London, Series A*. Article soumis.
- A. Arribas, M. Glover, A. Maidens, K. Peterson, M. Gordon, C. MacLachlan, R. Graham, D. Fereday, J. Camp, A. A. Scaife, P. Xavier, P. McLean, A. Colman, et S. Cusack (2011). The GloSea4 ensemble prediction system for seasonal forecasting. *Monthly Weather Review*, 139 : 1891–1910. doi : 10.1175/2010MWR3615.1.
- M. P. Baldwin, L. J. Gray, T. J. Dunkerton, K. Hamilton, P. H. Haynes, W. J. Randel, J. R. Holton, M. J. Alexander, I. Hirota, T. Horinouchi, D. B. A. Jones, J. S. Kinnersley, C. Marquardt, K. Sato, et M. Takahashi (2001). The Quasi-Biennial Oscillation. *Reviews of Geophysics*, 39(2) : 179–229.
- M. P. Baldwin, D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, et A. O'Neill (2003). Stratospheric memory and skill of extended-range weather forecasts. *Science*, 301 : 636–640.
- A. G. Barnston, M. H. Glantz, et Y. He (1999). Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. *Bulletin of the American Meteorological Society*, 80(2) : 217–244.
- M. Barreiro et P. Chang (2004). A linear tendency correction technique for improving seasonal prediction of SST. *Geophysical Research Letters*, 31 : L23209. doi : 10.1029/2004GL021148.
- L. Batté et M. Déqué (2011). Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models : skill of the ENSEMBLES project multimodel ensemble forecasts. *Tellus*, 63A : 283–299. doi : 10.1111/j.1600-0870.2010.00493.x.
- L. Batté et M. Déqué (2012). A stochastic method for improving seasonal predictions. *Geophysical Research Letters*, 39 : L09707. doi : 10.1029/2012GL051406.

- S. Belamari (2005). Report on uncertainty estimates of an optimal bulk formulation for surface turbulent fluxes. MERSEA IP deliverable. D.4.1.2, 29.
- L. Bengtsson, U. Schlese, E. Roeckner, M. Latif, T. P. Barnett, et N. E. Graham (1993). A two tiered approach to long-range climate forecasting. *Science*, 261 : 1026–1029.
- J. Berner, F. J. Doblas-Reyes, T. N. Palmer, G. Shutts, et A. Weisheimer (2008). Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. *Philosophical Transactions of the Royal Society of London, Series A*, 366(1875) : 2559–2577. doi : 10.1098/rsta.2008.0033.
- J. Berner, G. Shutts, M. Leutbecher, et T. N. Palmer (2009). A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependant predictability in the ECMWF Ensemble Prediction System. *Journal of the Atmospheric Sciences*, 66 : 603–626. doi : 10.1175/2008JAS2677.1.
- J. Berner, T. Jung, et T. N. Palmer (2012). Systematic model error : the impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *Journal of Climate*, 25 : 4946–4962. doi : 10.1175/JCLI-D-11-00297.1.
- L. Bouali (2009). *Prévisibilité et prévision statistico-dynamique des saisons des pluies associées à la mousson ouest africaine à partir d'ensembles multi-modèles*. PhD thesis, Université de Bourgogne, Ecole Doctorale “Environnement, Santé, STIC”.
- L. Bouali, N. Philippon, B. Fontaine, et J. Lemond (2008). Performance of DEMETER calibration for rainfall forecasting purposes : Application to the July – August Sahelian rainfall. *Journal of Geophysical Research*, 113(D15). doi : 10.1029/2007JD009403.
- P. Bougeault (1985). A simple parameterization of the large-scale effects of cumulus convection. *Monthly Weather Review*, 113 : 2108–2121.
- J. H. Bowden et F. H. M. Semazzi (2007). Empirical analysis of intraseasonal climate variability over the Greater Horn of Africa. *Journal of Climate*, 20(23) : 5715–5731.
- N. E. Bowler, A. Arribas, K. R. Mylne, K. B. Robertson, et S. E. Beare (2008). The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 134 : 703–722.
- C. Brankovic et T. N. Palmer (2000). Seasonal skill and predictability of ECMWF PROVOST ensembles. *Quarterly Journal of the Royal Meteorological Society*, 126(567) : 2035–2067.
- G. W. Brier (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78 : 1–3.
- R. Buizza, M. Miller, et T. N. Palmer (1999). Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society*, 125 : 2887–2908.
- P. Camberlin et N. Philippon (2002). The East-African March-May rainy season : Associated atmospheric dynamics and predictability over the 1968–1997 period. *Journal of Climate*, 15 (9) : 1002–1019.

- P. Camberlin, S. Janicot, et I. Pocard (2001). Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature : Atlantic VS ENSO. *International Journal of Climatology*, 21(8) : 973–1005.
- P. Camberlin, V. Moron, R. Okoola, N. Philippon, et W. Gitau (2009). Components of rainy seasons' variability in Equatorial East Africa : onset, cessation, rainfall frequency and intensity. *Theoretical and Applied Climatology*, 98 : 237–249. doi : 10.1007/s00704-009-0113-1.
- D. Cariolle et H. Teyssède (2007). A revised linear ozone photochemistry parameterization for use in transport and general circulation models : multi-annual simulations. *Atmospheric Chemistry and Physics*, 7 : 2183–2196.
- M. Charron, G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, et L. Michelin (2010). Toward random sampling of model error in the Canadian Ensemble Prediction System. *Monthly Weather Review*, 138 : 1877–1901.
- X. Cheng et J. M. Wallace (1993). Cluster analysis of Northern Hemisphere wintertime 500-hPa height field. *Journal of the Atmospheric Sciences*, 50 : 2674–2696.
- M. Chevallier et D. Salas y Mélia (2012). The role of sea ice thickness distribution in the Arctic sea ice potential predictability : a diagnostic approach with a coupled GCM. *Journal of Climate*, 25 : 3025–3038. doi : 10.1175/JCLI-D-11-00209.1.
- B. Christiansen (2007). Atmospheric circulation regimes : can cluster analysis provide the number ? *Journal of Climate*, 20 : 2229–2250.
- J. Cohen et D. Entekhabi (1999). Eurasian snow cover variability and Northern Hemisphere climate predictability. *Geophysical Research Letters*, 26 : 345–348.
- J. Cohen et C. Fletcher (2007). Improved skill of Northern Hemisphere winter surface temperature predictions based on land-atmosphere fall anomalies. *Journal of Climate*, 20 : 4118–4132. doi : 10.1175/JCLI4241.1.
- M. Collins, B. B. B. Booth, G. R. Harris, J. M. Murphy, D. M. H. Sexton, et M. J. Webb (2006). Towards quantifying uncertainty in transient climate change. *Climate Dynamics*, 27 : 127–147. doi : 10.1007/s00382-006-0121-0.
- S. Corti, S. Gualdi, et A. Navarra (2003). Analysis of the mid-latitude weather regimes in the 200-year control integration of the SINTEX model. *Annals of Geophysics*, 46 : 27–37.
- F. D'Andrea et R. Vautard (2000). Reducing systematic errors by empirically correcting model errors. *Tellus*, 52A : 21–41.
- C. M. Danforth, E. Kalnay, et T. Miyoshi (2007). Estimating and correcting global weather model error. *Monthly Weather Review*, 135 : 281–299. doi : 10.1175/MWR3289.1.
- D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, L. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, et al. (2011). The ERA-Interim reanalysis : configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137 : 553–597. doi : 10.1002/qj.828.

- P. Delecluse, M. K. Davey, Y. Kitamura, S. G. H. Philander, M. Suarez, et L. Bengtsson (1998). Coupled general circulation modeling of the tropical Pacific. *Journal of Geophysical Research*, 103 : 14357–14373. doi : 10.1029/97JC02546.
- T. DelSole, M. Zhao, P. A. Dirmeyer, et B. P. Kirtman (2008). Empirical correction of a coupled land-atmosphere model. *Monthly Weather Review*, 136 : 4063–4076. doi : 10.1175/2008MWR2344.1.
- M. Déqué (1997). Ensemble size for numerical seasonal forecasts. *Tellus*, 49A : 74–86.
- M. Déqué (2007). Estimation itérative de la tendance d’erreur dans ARPEGE. Note de centre, Centre National de Recherches Météorologiques, Météo-France.
- M. Déqué et J.-F. Royer (1992). The skill of extended-range extratropical winter dynamical forecasts. *Journal of Climate*, 5 : 1346–1356.
- M. Déqué, C. Dreveton, A. Braun, et D. Cariolle (1994). The ARPEGE/IFS atmosphere model : a contribution to the French community climate modelling. *Climate Dynamics*, 10 : 249–266.
- C. Deser (2000). On the teleconnectivity of the “Arctic Oscillation”. *Geophysical Research Letters*, 27(6) : 779–782.
- P. A. Dirmeyer (1999). Assessing GCM sensitivity to soil wetness using GSWP data. *Journal of the Meteorological Society of Japan*, 77 : 367–385.
- F. J. Doblas-Reyes, M. Déqué, et J.-P. Pielke (2000). Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Quarterly Journal of the Royal Meteorological Society*, 126 (567) : 2069–2087.
- F. J. Doblas-Reyes, V. Pavan, et D. B. Stephenson (2003). The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. *Climate Dynamics*, 21 : 501–514. doi : 10.1007/s00382-003-0350-4.
- F. J. Doblas-Reyes, R. Hagedorn, et T. N. Palmer (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus*, 57A : 234–252.
- F. J. Doblas-Reyes, A. Weisheimer, M. Déqué, N. Keenlyside, M. MacVean, J. M. Murphy, P. Rogel, D. Smith, et T. N. Palmer (2009). Addressing model uncertainty in seasonal and annual dynamical ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, 135 : 1538–1559. doi : 10.1002/qj.464.
- H. Douville (2010). Relative contribution of soil moisture and snow mass to seasonal climate predictability : a pilot study. *Climate Dynamics*, 34 : 797–818.
- H. Douville et F. Chauvin (2000). Relevance of soil moisture for seasonal climate predictions : a preliminary study. *Climate Dynamics*, 16 : 719–736.
- H. Douville, J.-F. Royer, et J.-F. Mahfouf (1995). A new snow parameterization for the Météo-France climate model. Part 1 : validation in stand-alone experiments. *Climate Dynamics*, 12 : 21–35.

- S. Engelstaedter et R. Washington (2007). Atmospheric controls on the annual cycle of North African dust. *Journal of Geophysical Research*, 112 : D03103. doi : 10.1029/2006JD007195.
- E. S. Epstein (1969a). Stochastic dynamic prediction. *Tellus*, 21 : 739–759.
- E. S. Epstein (1969b). A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology*, 8 : 985–987.
- B. Fontaine, N. Philippon, et P. Camberlin (1999). An improvement of June–September rainfall forecasting in the Sahel based upon region April–May moist static energy content (1968–1997). *Geophysical Research Letters*, 26(14) : 2041–2044.
- Y. Fouquart et B. Bonnel (1980). Computations of solar heating of the Earth’s atmosphere : a new parametrization. *Beitraege zur Physik der Atmosphaere*, 53 : 35–62.
- J. A. Francis, W. Chan, D. J. Leathers, J. R. Miller, et D. E. Veron (2009). Winter Northern Hemisphere weather patterns remember summer Arctic sea-ice extent. *Geophysical Research Letters*, 36(7) : L07503. doi : 10.1029/2009GL037274.
- G. Garric, H. Douville, et M. Déqué (2002). Prospects for improved seasonal predictions of monsoon precipitation over Sahel. *International Journal of Climatology*, 22 : 331–345. doi : 10.1002/joc.736.
- A. Giannini, R. Saravanan, et P. Chang (2003). Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science*, 302 : 1027–1030. doi : 10.1126/science.1089357.
- A. Giannini, R. Saravanan, et P. Chang (2005). Dynamics of the boreal summer African monsoon in the NSIPP1 atmospheric model. *Climate Dynamics*, 25 : 517–535. doi : 10.1007/s00382-005-0056-x.
- L. Goddard, S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, et M. A. Cane (2001). Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology*, 21 : 1111–1152. doi : 10.1002/joc.636.
- J.-F. Guérémy, M. Déqué, A. Braun, et J.-P. Piedelièvre (2005). Actual and potential skill of seasonal predictions using the CNRM contribution to DEMETER : coupled versus uncoupled model. *Tellus*, 57A : 308–319.
- A. Guldberg, E. Kaas, M. Déqué, S. Yang, et S. Vester Thorsen (2005). Reduction of systematic errors by empirical model correction : impact on seasonal prediction skill. *Tellus*, 57A : 575–588.
- R. Hagedorn, F. J. Doblas-Reyes, et T. N. Palmer (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting - I. Basic concept. *Tellus*, 57A : 219–233.
- M. Hamatan, G. Mahe, E. Servat, J.-E. Paturel, et A. Amani (2004). Synthèse et évaluation des prévisions saisonnières en Afrique de l’Ouest. *Sécheresse*, 15(3) : 279–286.
- A. R. Hansen et A. Sutera (1990). Weather regimes in a General Circulation Model. *Journal of the Atmospheric Sciences*, 47(3) : 380–391.
- K. Hasselmann (1976). Stochastic climate models. Part I : Theory. *Tellus*, 28 : 473–485.

- R. N. Hoffman et E. Kalnay (1983). Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, 35A : 100–118.
- M. Honda, J. Inoue, et S. Yamane (2009). Influence of low Arctic sea-ice minima on anomalously cold Eurasian winters. *Geophysical Research Letters*, 36(8) : L08707. doi : 10.1029/2008GL037079.
- J. D. Horel et J. M. Wallace (1981). Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Monthly Weather Review*, 109 : 813–829.
- M. Hortal et A. J. Simmons (1991). Use of reduced Gaussian grids in spectral models. *Monthly Weather Review*, 119 : 1057–1074.
- B. Hoskins et P. S. Schopf (2008). Ocean-atmosphere basis for seasonal climate forecasting. In A. Troccoli, M. Harrison, D. L. T. Anderson, et S. J. Mason, editors, *Seasonal Climate : Forecasting and Managing Risk*, pages 67–89. Springer Academic Publishers : Dordrecht, Netherlands.
- J. Huang, H. M. van den Dool, et A. G. Barnston (1996). Long-lead seasonal temperature prediction using Optimal Climate Normals. *Journal of Climate*, 9 : 809–817.
- G. J. Huffman, R. F. Adler, D. T. Bolvin, et G. Gu (2009). Improving the global precipitation record : GPCP version 2.1. *Geophysical Research Letters*, 36. ISSN 0094-8276. doi : 10.1029/2009GL040000.
- S. Janicot, S. Trzaska, et I. Poccarrd (2001). Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dynamics*, 18 : 303–320.
- A. B. M. Jeuken, P. C. Siegmund, L. C. Heijboer, J. Feichter, et L. Bengtsson (1996). On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *Journal of Geophysical Research*, 101 : 16939–16950.
- M. Joly et A. Voldoire (2009). Influence of ENSO on the West African Monsoon : Temporal aspects and atmospheric processes. *Journal of Climate*, 22 : 3193–3210. doi : 10.1175/2008JCLI2450.1.
- T. Jung, T. N. Palmer, et G. J. Shutts (2005). Influence of a stochastic parameterization on the frequency of occurrence of north pacific weather regimes in the ECMWF model. *Geophysical Research Letters*, 32 : L23811. doi : 10.1029/2005GL024248.
- T. Jung, M. J. Miller, T. N. Palmer, P. Towers, N. Wedi, D. Achuthavarier, J. M. Adams, E. L. Altshuler, B. A. Cash, J. L. Klinger III, L. Marx, C. Stan, et K. I. Hodges (2012). High-resolution global climate simulations with the ECMWF model in Project Athena : Experimental design, model climate, and seasonal forecast skill. *Journal of Climate*, 25 : 3155–3172. doi : 10.1175/JCLI-D-11-00265.1.
- E. Kaas, A. Guldberg, W. May, et M. Déqué (1999). Using tendency errors to tune the parameterisation of unresolved dynamical scale interactions in atmospheric general circulation models. *Tellus*, 51A : 612–629.
- E. Kalnay (2003a). Data assimilation. In *Atmospheric Modelling, Data Assimilation and Predictability*, chapter 5. Cambridge University Press.

- E. Kalnay (2003b). Atmospheric predictability and ensemble forecasting. In *Atmospheric Modelling, Data Assimilation and Predictability*, chapter 6. Cambridge University Press.
- M. Kimoto et M. Ghil (1993a). Multiple flow regimes in the Northern Hemisphere winter. Part I : Methodology and hemispheric regimes. *Journal of the Atmospheric Sciences*, 50 : 2625–2643.
- M. Kimoto et M. Ghil (1993b). Multiple flow regimes in the Northern Hemisphere winter. Part II : Sectorial regimes and preferred transitions. *Journal of the Atmospheric Sciences*, 50 : 2645–2672.
- B. P. Kirtman, J. Shukla, B. Huang, Z. Zhu, et E. K. Schneider (1997). Multiseasonal predictions with a coupled Tropical Ocean-Global Atmosphere system. *Monthly Weather Review*, 125 : 789–808.
- R. D. Koster, P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanane, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, et T. Yamada (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, 305 : 1138–1140. doi : 10.1126/science.1100217.
- R. D. Koster, S. P. P. Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, P. A. Dirmeyer, F. J. Doblas-Reyes, G. Drewitt, C. T. Gordon, Z. Guo, J.-H. Jeong, D. M. Lawrence, W.-S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B. J. J. M. van den Hurk, F. Vitart, et E. F. Wood (2010). Contribution of land surface initialization to subseasonal forecast skill : First results from a multi-model experiment. *Geophysical Research Letters*, 37(L02402). doi : 10.1029/2009GL041677.
- T. N. Krishnamurti, C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, et S. Surendran (1999). Improved weather and seasonal climate forecasts from multimodel superensembles. *Science*, 285(5433) : 1548–1550.
- C. E. Leith (1974). Theoretical skill of Monte Carlo forecasts. *Monthly Weather Review*, 102 : 409–418.
- R. E. Livezey, K. Y. Vinnikov, M. M. Timofeyeva, R. Tinker, et H. M. van den Dool (2007). Estimation and extrapolation of climate normals and climatic trends. *Journal of Applied Meteorology and Climatology*, 46 : 1759–1776.
- E. N. Lorenz (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20 : 130–141.
- E. N. Lorenz (1969). The predictability of a flow which possesses many scales of motion. *Tellus*, 21 : 289–307.
- T. Losada, B. Rodriguez-Fonseca, E. Mohino, J. Bader, S. Janicot, et C. R. Mechoso (2012). Tropical SST and Sahel rainfall : a non-stationary relationship. *Geophysical Research Letters*, 39 : L12705. doi : 10.1029/2012GL052423.
- J.-F. Louis (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Meteorology*, 17 : 187–202.

- J. MacQueen (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics and Probabilities*, volume 1, pages 281–297. University of California Press.
- G. Madec (2008). NEMO ocean engine. Note du Pôle de modélisation No 27, Institut Pierre-Simon Laplace (IPSL), France. ISSN No 1288-1619.
- A. J. Majda, C. Franzke, et B. Khouider (2008). An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society of London, Series A*, 366 : 2427–2453. doi : 10.1098/rsta.2008.0012.
- S. Manabe, J. Smagorinsky, et R. F. Strickler (1965). Simulated climatology of a general circulation model with a hydrologic cycle. *Monthly Weather Review*, 93(12) : 769–798.
- J. Marshall et F. Molteni (1993). Toward a dynamical understanding of planetary-scale flow regimes. *Journal of the Atmospheric Sciences*, 50 : 1792–1818.
- S. J. Mason (2008). From dynamical model predictions to seasonal climate forecasts. In A. Troccoli, M. Harrison, D. L. T. Anderson, et S. J. Mason, editors, *Seasonal Climate : Forecasting and Managing Risk*, pages 205–234. Springer Academic Publishers : Dordrecht, Netherlands.
- S. J. Mason et O. Baddour (2008). Statistical modelling. In A. Troccoli, M. Harrison, D. L. T. Anderson, et S. J. Mason, editors, *Seasonal Climate : Forecasting and Managing Risk*, pages 163–201. Springer Academic Publishers : Dordrecht, Netherlands.
- P.-A. Michelangeli, R. Vautard, et B. Legras (1995). Weather Regimes : Recurrence and Quasi Stationarity. *Journal of the Atmospheric Sciences*, 52(8) : 1237–1256.
- E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, et S. A. Clough (1997). Radiative transfer for inhomogeneous atmospheres : RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, 102 : 16663–16682.
- K. C. Mo et W. M. Thiaw (2002). Ensemble canonical correlation prediction of precipitation over the Sahel. *Geophysical Research Letters*, 29(12). doi : 10.1029/2002GL015075.
- F. Molteni (2003). Atmospheric simulations using a GCM with simplified physical parametrizations. I : Model climatology and variability in multi-decadal experiments. *Climate Dynamics*, 20 : 175–191.
- A. H. Murphy (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, 12 : 595–600.
- J. M. Murphy, D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, et D. A. Stainforth (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430 : 768–772.
- O. Ndiaye, M. N. Ward, et W. M. Thiaw (2011). Predictability of seasonal Sahel rainfall using GCMs and lead-time improvements through the use of a coupled model. *Journal of Climate*, 24 : 1931–1949.
- J. D. Neelin, D. S. Battisti, A. C. Hirst, J. Fei-Fei, Y. Wakata, T. Yamagata, et S. Zebiak (1998). ENSO theory. *Journal of Geophysical Research*, 103 : 14261–14290.

- J. Noilhan et J.-F. Mahfouf (1996). The ISBA land surface parameterisation scheme. *Global Planet Change*, 13 : 145–159.
- J. Noilhan et S. Planton (1989). A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, 117 : 536–549.
- T. Oki et Y. C. Sud (1998). Design of total runoff integrating pathways (TRIP). a global river channel network. *Earth Interactions*, 2 : 1–36.
- Y. J. Orsolini, R. Senan, R. E. Benestad, et A. Melsom (2012). Autumn atmospheric response to the 2007 low Arctic sea ice extent in coupled ocean–atmosphere hindcasts. *Climate Dynamics*, 38(11) : 2437–2448.
- G. Ouzeau, J. Cattiaux, H. Douville, A. Ribes, et D. Saint-Martin (2011). European cold winter 2009-2010 : How unusual in the instrumental record and how reproducible in the ARPEGE-Climat model? *Geophysical Research Letters*, 38(L11706). doi : 10.1029/2011GL047667.
- T. N. Palmer (2001). A nonlinear dynamical perspective on model error : A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, 127 : 279–304.
- T. N. Palmer (2002). The economic value of ensemble forecasts as a tool for risk assessment : from days to decades. *Quarterly Journal of the Royal Meteorological Society*, 128 : 747–774.
- T. N. Palmer (2009). Edward Norton Lorenz. *Biographical Memoirs of Fellows of the Royal Society*, 55 : 139–155.
- T. N. Palmer (2012). Towards the probabilistic Earth-system simulator : a vision for the future of climate and weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 138 : 841–861. doi : 10.1002/qj.1923.
- T. N. Palmer et J. Shukla (2000). Editorial, DSP/PROVOST Special Issue. *Quarterly Journal of the Royal Meteorological Society*, 126(567).
- T. N. Palmer, A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Décluse, M. Déqué, E. Díez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. S. Keenlyside, M. Latif, A. Lazar, E. Maisonave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J.-M. Terres, et M. C. Thomson (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society*, 85 : 853–872.
- T. N. Palmer, R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, et L. Smith (2006). Ensemble prediction : a pedagogical perspective. In *ECMWF Newsletter*, volume 106, pages 10–17.
- T. N. Palmer, R. Buizza, F. Doblas-Reyes, F. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, et A. Weisheimer (2009). Stochastic parametrization and model uncertainty. Technical Memorandum 598, ECMWF.
- Y. Peings et H. Douville (2009). Influence of the Eurasian snow cover on the Indian summer monsoon variability in observations and CMIP3 simulations. *Climate Dynamics*, 34 : 643–660. doi : 10.1007/s00382-009-0565-0.

- C. Penland (2003). Noise out of chaos and why it won't go away. *Bulletin of the American Meteorological Society*, 84 : 921–925. doi : 10.1175/BAMS-84-7-921.
- C. Penland et B. D. Ewald (2008). On modelling physical systems with stochastic models : diffusion versus Lévy processes. *Philosophical Transactions of the Royal Society of London, Series A*, 366 : 2455–2474. doi : 10.1098/rsta.2008.0051.
- N. Philippon, F. J. Doblas-Reyes, et P. M. Ruti (2010). Skill, reproducibility and potential predictability of the West African monsoon in coupled GCMs. *Climate Dynamics*, 35 : 53–74. doi : 10.1007/s00382-010-0856-5.
- M. Rajeevan, D. S. Pai, R. A. Kumar, et B. Lal (2007). New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics*, 28 : 813–828. doi : 10.1007/s00382-006-0197-6.
- J.-L. Ricard et J.-F. Royer (1993). A statistical cloud scheme for use in an AGCM. *Annales Geophysicae*, 11 : 1095–1115.
- D. S. Richardson (2003). Economic value and skill. In I. Joliffe et D. Stephenson, editors, *Forecast Verification : A Practitioner's Guide in Atmospheric Science*, pages 165–187. John Wiley & Sons Ltd.
- M. M. Rienecker, M. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W. Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, et al. (2011). MERRA-NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*. doi : 10.1175/JCLI-D-11-00015.1.
- P. Rogel, Y. M. Turre, V. Benoit, et L. Jarlan (2006). Tropical Atlantic moisture availability and precipitation over West Africa : Application to DEMETER hindcasts. *Geophysical Research Letters*, 33. doi : 10.1029/2006GL027178.
- M. Rouault et Y. Richard (2005). Intensity and spatial extent of droughts in southern Africa. *Geophysical Research Letters*, 32. doi : 10.1029/2005GL022436.
- J.-F. Royer, S. Planton, et M. Déqué (1990). A sensitivity experiment for the removal of Arctic sea ice with the French spectral general circulation model. *Climate Dynamics*, 5(1) : 1–17.
- S. Saha, S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, Y.-T. Hou, H.-Y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, et al. (2010). The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, 91 : 1015–1057.
- D. Salas y Melia (2002). A global coupled sea ice–ocean model. *Ocean Modeling*, 4 : 137–172. doi : 10.1016/S1463-5003(01)00015-4.
- R. Sausen, K. Barthel, et K. Hasselmann (1988). Coupled ocean-atmosphere models with flux correction. *Climate Dynamics*, 2 : 145–163.
- U. Schneider, T. Fuchs, A. Meyer-Christoffer, et B. Rudolf (2008). Global precipitation analysis products of the GPCC. *Global Precipitation Climatology Centre (GPCC) Internet Publication*, pages 1–12.

- J. Shukla (1981). Dynamical predictability of monthly means. *Journal of the Atmospheric Sciences*, 38(12) : 2547–2572.
- J. Shukla (1998). Predictability in the midst of chaos : a scientific basis for climate forecasting. *Science*, 282 : 728–731. doi : 10.1126/science.282.5389.728.
- G. Shutts (2005). A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 131 : 3079–3102. doi : 10.1256/qj.04.106.
- D. B. Stephenson (2008). An introduction to probability forecasting. In A. Troccoli, M. Harrison, D. L. T. Anderson, et S. J. Mason, editors, *Seasonal Climate : Forecasting and Managing Risk*, pages 235–257. Springer Academic Publishers : Dordrecht, Netherlands.
- D. B. Stephenson, A. Hannachi, et A. O’Neill (2004). On the existence of multiple climate regimes. *Quarterly Journal of the Royal Meteorological Society*, 130 : 583–606.
- D. B. Stephenson, C. A. S. Coelho, F. J. Doblas-Reyes, et M. Balmaseda (2005). Forecast assimilation : a unified framework for the combination of multi-model weather and climate predictions. *Tellus*, 57A(3) : 253–264.
- B. Sultan, S. Janicot, et A. Diedhiou (2003). The West African Monsoon dynamics. Part I : Documentation of intraseasonal variability. *Journal of Climate*, 16(21) : 3389–3406.
- S. Szopa, Y. Balkanski, M. Schulz, S. Bekki, D. Cugnet, A. Fortems-Cheiney, S. Turquety, A. Cozic, C. Déandreis, D. Hauglustaine, A. Idelkadi, J. Lathière, F. Lefevre, M. Marchand, R. Vuolo, N. Yan, et J.-L. Dufresne (2012). Aerosol and ozone changes as forcing for climate evolution between 1850 and 2100. *Climate Dynamics*. doi : 10.1007/s00382-012-1408-y.
- D. W. J. Thompson et J. M. Wallace (2000). Annular modes in extratropical circulation. Part I : Month-to-month variability. *Journal of Climate*, 13 : 1000–1016.
- Z. Toth, O. Talagrand, G. Candille, et Y. Zhu (2003). Probability and ensemble forecasts. In I. Joliffe et D. Stephenson, editors, *Forecast Verification, A Practitioner’s Guide in Atmospheric Science*, pages 137–163. John Wiley & Sons Ltd.
- K. E. Trenberth (1997). The definition of El Niño. *Bulletin of the American Meteorological Society*, 78 : 2771–2777.
- K. E. Trenberth, G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, et C. Ropelewski (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research*, 103 : 14291–14324.
- A. Troccoli (2010). Review : Seasonal climate forecasting. *Meteorological Applications*, 17 : 251–268. doi : 10.1002/met.184.
- UNEP (2010). Africa Water Atlas. Division of Early Warning and Assessment, United Nations Environment Programme.
- S. M. Uppala, P. W. Kallberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De Berg, et al. (2005). The ERA-40 reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 131 : 2961–3012. doi : 10.1256/qj.04.176.

- S. Valcke (2006). OASIS3 user guide. Technical Report TR/CMGC/06/73, CERFACS, Toulouse, France.
- B. van den Hurk, F. Doblas-Reyes, G. Balsamo, R. D. Koster, S. I. Seneviratne, et H. Camargo Jr (2012). Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dynamics*, 38 : 349–362. doi : 10.1007/s00382-010-0956-2.
- H. van Loon et K. Labitzke (1987). The Southern Oscillation. Part V : The anomalies in the lower stratosphere of the Northern Hemisphere in winter and a comparison with the quasi-biennial oscillation. *Monthly Weather Review*, 115 : 357–369.
- G. J. van Oldenborgh, M. A. Balmaseda, L. Ferranti, T. N. Stockdale, et D. L. T. Anderson (2005). Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? *Journal of Climate*, 18 : 3240–3249.
- E. K. Vizy et K. H. Cook (2001). Mechanisms by which Gulf of Guinea and Eastern North Atlantic sea surface temperature anomalies can influence African Rainfall. *Journal of Climate*, 14(5) : 795–821.
- A. Voldoire, E. Sanchez-Gomez, D. Salas y Mélia, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, M. Chevallier, M. Déqué, J. Deshayes, H. Douville, E. Fernandez, G. Madec, E. Maisonnave, M.-P. Moine, S. Planton, D. Saint-Martin, S. Szopa, et al. (2012). The CNRM-CM5.1 global climate model : Description and basic evaluation. *Climate Dynamics*. doi : 10.1007/s00382-011-1259-y.
- G. T. Walker et E. W. Bliss (1932). World weather V. *Memoirs of the Royal Meteorological Society*, 4 : 53–84.
- J. M. Wallace et D. S. Gutzler (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review*, 109 : 784–812.
- B. Wang, J.-Y. Lee, I.-S. Kang, J. Shukla, C.-K. Park, A. Kumar, J. Schemm, S. Cocke, J.-S. Kug, J.-J. Luo, L. Zhou, B. Wang, X. Fu, W.-T. Yun, O. Alves, E. K. Jin, J. Kinter, B. Kirtman, T. Krishnamurti, N. C. Lau, et al. (2009). Advance and prospectus of seasonal prediction : assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dynamics*, 33 : 93–117. doi : 10.1007/s00382-008-0460-0.
- A. P. Weigel, M. A. Liniger, et C. Appenzeller (2008). Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society*, 134 : 241–260. doi : 10.1002/qj.210.
- A. Weisheimer, F. J. Doblas-Reyes, T. N. Palmer, A. Alessandri, A. Arribas, M. Déqué, N. Keenlyside, M. MacVean, A. Navarra, et P. Rogel (2009). ENSEMBLES : A new multi-model ensemble for seasonal-to-annual predictions — skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophysical Research Letters*, 36. doi : 10.1029/2009GL040896.
- A. Weisheimer, T. N. Palmer, et F. J. Doblas-Reyes (2011). Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophysical Research Letters*, 38 : L16703. doi : 10.1029/2011GL048123.
- D. S. Wilks (2006). Forecast Verification. In *Statistical Methods in the Atmospheric Sciences, Second Edition*, chapter 7, pages 255–335. Academic Press.

- D. S. Wilks (2008). Effects of stochastic parametrization on conceptual climate models. *Philosophical Transactions of the Royal Society of London, Series A*, 366 : 2477–2490. doi : 10.1098/rsta.2008.0005.
- P. D. Williams (2012). Climatic impacts of stochastic fluctuations in air-sea fluxes. *Geophysical Research Letters*, 39 : L10705. doi : 10.1029/2012GL051813.
- X.-Q. Yang et J. L. Anderson (2000). Correction of systematic errors in coupled GCM forecasts. *Journal of Climate*, 13 : 2072–2085.
- S. E. Zebiak et M. A. Cane (1987). A model El Niño-Southern Oscillation. *Monthly Weather Review*, 115 : 2262–2278.